Australian Government
Department of Defence
Defence Science and
Technology Organisation

A Review of Accident Modelling Approaches for Complex Critical Sociotechnical Systems

Zahid H. Qureshi

Command, Control, Communications and Intelligence Division[1]
Defence Science and Technology Organisation

DSTO-TR-2094

ABSTRACT

The increasing complexity in highly technological systems such as aviation, maritime, air traffic control, telecommunications, nuclear power plants, defence and aerospace, chemical and petroleum industry, and healthcare and patient safety is leading to potentially disastrous failure modes and new kinds of safety issues. Traditional accident modelling approaches are not adequate to analyse accidents that occur in modern sociotechnical systems, where accident causation is not the result of an individual component failure or human error. This report provides a review of key traditional accident modelling approaches and their limitations, and describes new system-theoretic approaches to the modelling and analysis of accidents in safety-critical systems. It also discusses current research on the application of formal (mathematically-based) methods to accident modelling and organisational theories on safety and accident causation. This report recommends new approaches to the modelling and analysis of complex systems that are based on systems theory and interdisciplinary research, in order to capture the complexity of modern sociotechnical systems from a broad systemic view for understanding the multi-dimensional aspects of safety and accident causation.

RELEASE LIMITATION

Approved for public release

[1] Current Address: Defence and Systems Institute, Division of Information Technology, Engineering and the Environment, University of South Australia.

Published by

Command, Control, Communications and Intelligence Division
DSTO Defence Science and Technology Organisation
PO Box 1500
Edinburgh South Australia 5111 Australia

Telephone: (08) 8259 5555
Fax: (08) 8259 6567

© Commonwealth of Australia 2008
AR-014-089
January 2008

APPROVED FOR PUBLIC RELEASE

A Review of Accident Modelling Approaches for Complex Critical Sociotechnical Systems

Executive Summary

Highly technological systems such as aviation, maritime, air traffic control, telecommunications, nuclear power plants, defence and aerospace, chemical and petroleum industry, and healthcare and patient safety are exceedingly becoming more complex. Such complex systems can exhibit potentially disastrous failure modes. Notable disasters and accidents such as the Bhopal toxic gas release disaster (Srivastava, 1992), the NASA Challenger shuttle explosion (Vaughn, 1996), the US Black Hawk fratricide incident during the 1994 Gulf War Operation Provide Comfort (AAIB, 1994), the Royal Australian Air Force F-111 chemical exposure of maintenance workers (Clarkson et al., 2001), the Esso Longford gas plant accident (Hopkins, 2000), and a number of critical aviation and train accidents such as the 1993 Warsaw accident (Höhl & Ladkin, 1997) and the Glenbrook NSW Rail accident (Ladkin, 2005) respectively, are clear examples of system failures in complex systems that led to serious loss of material and human life.

Large complex systems such as the Bhopal chemical plant and the Operation Provide Comfort Command and Control System are semantically complex (it generally takes a great deal of time to master the relevant domain knowledge), with tight couplings between various parts, and where operations are often carried out under time pressure or other resource constraints (Woods et al., 1994). In such systems, accidents gradually develop over a period of time through a conjunction of several small failures, both machine and human (Perrow, 1984; Reason, 1990). This pattern is generally found in different industrial and aerospace accidents, despite the fact that every sociotechnical system is unique and each accident has many different aspects.

It is important to understand the causes of accidents in complex systems in order to enhance the safety of such systems, and to develop preventative strategies to mitigate the occurrence of future similar accidents. Accident models provide a conceptualisation of the characteristics of the accident, which typically show the relation between causes and effects. They explain why accidents occur, and are used as techniques for risk assessment during system development, and for *post hoc* accident analysis to study the causes of the occurrence of an accident. Most of the engineering models originated before the introduction of digital technology; these models have been updated but have not kept pace with the fast change in technological revolution. Modern technology is having a significant impact on the nature of accidents, and this requires new causal explanatory mechanisms to understand them and in the development of new risk assessment techniques to prevent their occurrence (Leveson, 2003).

Traditionally, accidents have been viewed as resulting from a chain of failure events, each related to its "causal" event or events. Almost all safety analysis and risk assessment techniques are based on this linear notion of causality, which have severe

limitations in the modelling and analysis of modern complex systems. As opposed to conventional engineered systems, modern complex systems constitute different kinds of elements, intentional and non-intentional: social institutions, human agents and technical artefacts (Kroes et al., 2006). In these systems, referred as sociotechnical systems, humans interact with technology to deliver outcomes that cannot be attained by humans or technology functioning in isolation. In sociotechnical systems human agents and social institutions are integrated, and the attainment of organisational objectives is not met by the optimisation of technical systems alone, but by the joint optimisation of the technical and social aspects (Trist & Bamforth, 1951). Thus, the study of modern complex systems requires an understanding of the interactions and interrelationships between the technical, human, social and organisational aspects of systems. These interactions and interrelationships are complex and non-linear, and traditional modelling approaches cannot fully analyse the behaviours and failure modes of such systems.

In this report, we provide a review of key traditional accident modelling approaches and their limitations in capturing accident causality and dynamics of modern complex systems. We discuss new approaches to safety and accident modelling of sociotechnical systems that are based on systems theory and cognitive systems engineering. Systems theory includes the principles, models, and laws necessary to understand complex interrelationships and interdependencies between components (technical, human, organisational and management) of a complex system. Cognitive systems engineering (Hollnagel & Woods, 1983) provides a framework to model the behaviour of joint human-machine systems in the context of the environment in which work takes place. We also review the current research in formal (mathematically-based) methods for the modelling of complex system accidents. In addition, organisational sociologists have made significant contributions to the understanding of accidents in complex sociotechnical systems. Vaughn (1996) rejects the prevalent explanations (provided by traditional safety engineering techniques) of the cause of the *Challenger* shuttle accident and presents an alternative sociological explanation that explores much deeper cause of the failure.

The findings of this survey recommend new approaches to the modelling and analysis of complex systems that are based on systems theory. The sociotechnical system must be treated as an integrated whole, and the emphasis should be on the simultaneous consideration of social and technical aspects of systems, including social structures and cultures, social interaction processes, and individual factors such as capability and motivation as well as engineering design and technical aspects of systems. Interdisciplinary research is needed to capture the complexity of modern sociotechnical systems from a broad systemic view for understanding the multi-dimensional aspects of safety and modelling sociotechnical system accidents.

Author

Zahid H. Qureshi
Defence and Systems Institute, Division of Information Technology, Engineering and the Environment, University of South Australia

Zahid Qureshi graduated with a B.Sc. (Hons.) in Engineering (Electronics & Telecommunications) from the University of Engineering and Technology, Lahore, Pakistan in 1975. He was awarded a University of Wollongong post-graduate research award where he undertook a Ph.D. in Systems Science completing it in 1983. He has worked in industry in Asia and Australia for over 10 years on many systems and software engineering projects, including real-time software development, rail transportation, and formal specification and verification of safety-critical software. He held a faculty position with the Nanyang Technological University in Singapore from 1991 to 1994. Zahid joined DSTO as a senior research scientist in January 1996, and spent 10 years in the Air Operations Division, Information Technology Division and the Command & Control Division. His major projects included research in Aviation Automation and leading a long range research task on complex systems modelling and analysis of avionics mission systems. Zahid also investigated new approaches to accident modelling of complex critical sociotechnical systems based on systems theory and cognitive systems engineering. He is now a senior lecturer in the Defence and Systems Institute at the University of South Australia. His current research interests are focussed on a multidisciplinary approach to the modelling and analysis of complex sociotechnical system safety and accident causation, which considers not just the technical aspects but also the human, social, cultural and organisational factors.

Contents

1. Introduction

System safety is generally considered as the characteristics of a system that prevents injury to or loss of human life, damage to property, and adverse consequences to the environment. The IEC 61508 (1998-2000) safety standard defines safety as, "freedom from unacceptable risk of physical injury or of damage to the health of people, either directly, or indirectly as a result of damage to property or to the environment".

Highly technological systems such as aviation, maritime, air traffic control, telecommunications, nuclear power plants, defence and aerospace, chemical and petroleum industry, and healthcare and patient safety are exceedingly becoming more complex. Such complex systems can exhibit potentially disastrous failure modes. Notable disasters and accidents such as the Bhopal toxic gas release disaster (Srivastava, 1992), the NASA Challenger shuttle explosion (Vaughn, 1996), the US Black Hawk fratricide incident during the 1994 Gulf War Operation Provide Comfort (AAIB, 1994), the Royal Australian Air Force F-111 chemical exposure of maintenance workers (Clarkson et al., 2001), the Esso Longford gas plant accident (Hopkins, 2000), and a number of critical aviation and train accidents such as the 1993 Warsaw accident (Höhl & Ladkin, 1997) and the Glenbrook NSW Rail accident (Ladkin, 2005) respectively, are clear examples of system failures in complex systems that led to serious loss of material and human life.

Bhopal is the site of probably the greatest industrial disaster in history. In the early hours of 3rd December 1984, a pesticide plant owned by Union Carbide, a US-based multinational company, released a cloud of deadly gas into the atmosphere (Srivastava, 1992). Within minutes, it had drifted over the sleeping town of Bhopal in India. Estimates of the number of deaths on that night vary widely. The Indian government's official estimate is that 1,700 people died within 48 hours. Unofficially, it is said that around 6,000 people perished in the days immediately following the gas leak. What is certain is that the victims of Bhopal suffered horribly, most of them drowning in their own bodily fluids as the gas attacked their lungs. To date, over 20,000 people have died as a result of the accident. An estimated 10-15 people suffer crippling, gas-related deaths each month. More than 50,000 are too sick to work, while around 5,000 families continue to drink poisoned water. As a result, the infant mortality rate is significantly higher in Bhopal than in the rest of the country. The Bhopal disaster was a result of a combination of legal, technological, organisational, and human errors (Rasmussen, 1997).

One of the worst air-to-air friendly fire accidents involving US aircraft in military history occurred on April 14, 1994 over northern Iraq (AAIB, 1994) during Operation Provide Comfort. A pair of F-15Cs of the 52nd Fighter Wing enforcing the No Fly Zone mistakenly shot down two UH-60 Black Hawk helicopters, killing 26 American and United Nations personnel who were carrying out humanitarian aid to Kurdish areas of Iraq. One of the helicopters was destroyed by an AIM-120, the other by a Sidewinder. After a series of investigations by military and civilian boards with virtually unlimited resources, no culprit emerged; no bad guy showed himself and no smoking gun was found (Snook, 2002). The major reasons for what went wrong were organisational factors and the human operational use of technical systems that were embedded in a complex Command and Control structure (Leveson et al., 2002). Furthermore, it should be noted that, except for the failure of the Identify Friend or Foe (IFF) equipment, there were no technical malfunctions which contributed to the accident.

Large complex systems such as the Bhopal chemical plant and the Operation Provide Comfort Command and Control System are semantically complex (it generally takes a great deal of time to master the relevant domain knowledge), with tight couplings between various parts, and where operations are often carried out under time pressure or other resource constraints (Woods et al., 1994). In such systems, accidents gradually develop over a period of time through a conjunction of several small failures, both machine and human (Perrow, 1984; Reason, 1990). This pattern is generally found in different industrial and aerospace accidents, despite the fact that every sociotechnical system is unique and each accident has many different aspects.

Accident models provide a conceptualisation of the characteristics of the accident, which typically show the relation between causes and effects. They explain why accidents occur, and are used as techniques for risk assessment during system development, and for *post hoc* accident analysis to study the causes of the occurrence of an accident. Most of the engineering models originated before the introduction of digital technology; these models have been updated but have not kept pace with the fast change in technological revolution. Modern technology is having a significant impact on the nature of accidents, and this requires new causal explanatory mechanisms to understand them and in the development of new risk assessment techniques to prevent their occurrence (Leveson, 2003).

The historical development of accident models and various approaches for accident analysis have been discussed by engineers, scientists, cognitive psychologists, and sociologists (Ferry, 1988; Hayhurst & Holloway, 2003; Hollnagel & Woods, 2005; Johnson, 2003; Leveson, 1995; Leveson, 2001; Perrow, 1984; Rasmussen & Svedung, 2000; Reason, 1997; Skelt, 2002; Vaughn, 1996). In particular, Hollnagel (2001) provides an overview of the major changes to accident models since the 1950s, and argues that this reflects the developments in the commonly agreed understandings of the nature of an accident.

One of the earliest accident causation models is the Domino theory proposed by Heinrich in the 1940s (Heinrich et al., 1980), which describes an accident as a chain of discrete events which occur in a particular temporal order. This theory belongs to the class of sequential accident models or event-based accident models, which underlie most accident models such as Failure Modes and Effects Analysis (FMEA), Fault Tree Analysis (FTA), Event Tree Analysis, and Cause-Consequence Analysis (Leveson, 1995). These models work well for losses caused by failures of physical components or human errors in relatively simple systems. Typically, in these models, causal factors in an accident which was not linked to technical component failures were classified as human error as a kind of catchall or "garbage can" (Hollnagel, 2001). These models are limited in their capability to explain accident causation in the more complex systems that were developed in the last half of the 20th century.

In the 1980s, a new class of epidemiological accident models endeavoured to explain accident causation in complex systems. Epidemiological models regard events leading to accidents as analogous to the spreading of a disease, i.e., as the outcome of a combination of factors, some manifest and some latent, that happen to exist together in space and time. Reason's (1990; 1997) Swiss cheese model of defences is a major contribution to this class of models, and has greatly influenced the understanding of

accidents by highlighting the relationship between latent and immediate causes of accidents.

Sequential and epidemiological accident models are inadequate to capture the dynamics and nonlinear interactions between system components in complex sociotechnical systems. New accident models, based on systems theory, classified as systemic accident models, endeavour to describe the characteristic performance on the level of the system as a whole, rather than on the level of specific cause-effect "mechanisms" or even epidemiological factors (Hollnagel, 2004). Systemic safety models have their roots in systems theory and cybernetics. Systems theory includes the principles, models, and laws necessary to understand complex interrelationships and interdependencies between components (technical, human, organisational and management) of a complex system. Safety models based on systems theory view accidents as emergent phenomena that arise from interactions among system components, where the interactions may be nonlinear and involve multiple feedback loops (Perrow, 1984).

A major difference between systemic accident models and sequential/epidemiological accident models is that systemic accident models describe an accident process as a complex and interconnected network of events while the latter describes it as a simple cause-effect chain of events. Two notable systemic modelling approaches, Rasmussen's (1997) hierarchical sociotechnical framework and Leveson's (2004) STAMP (Systems-Theoretic Accident Model and Processes) model, endeavour to model the dynamics of complex sociotechnical systems.

Modern technology and automation has significantly changed the nature of human work from mainly manual tasks to predominantly knowledge intensive activities and cognitive tasks. This has created new problems for human operator performance (such as cognitive load) and new kinds of failure modes in the overall human-machine systems. Cognitive systems engineering (Hollnagel & Woods, 1983) has emerged as a framework to model the behaviour of human-machine systems in the context of the environment in which work takes place. Two systemic accident models for safety and accident analysis have been developed based on the principles of cognitive systems engineering: CREAM - Cognitive Reliability and Error Analysis Method (Hollnagel, 1998); and FRAM - Functional Resonance Accident Model (Hollnagel, 2004).

During the last decade many attempts have been made on the use of formal methods for building mathematically-based models to conduct accident analysis (Fields et al., 1995; Burns, 2000; Johnson & Holloway, 2003a; Vernez et al., 2003)). Formal methods can improve accident analysis by emphasising the importance of precision in definitions and descriptions, and providing notations for describing and reasoning about certain aspects of accidents. One of the most advanced application of formal methods to accident analysis is the *Why-Because Analysis* method (Ladkin & Loer, 1998), which employs a formal logic for accident modelling and rigorous reasoning for causal analysis. This method has been successfully applied to a number of case studies in aviation and rail transportation (Höhl & Ladkin, 1997; Ladkin, 2005).

As the understanding of industrial, transportation and aerospace accidents has evolved, they are no longer considered as simply the failures of technology alone, nor solely arising from the ubiquitous "human error", but also as a result of a historical background and an unfavourable organisational context (Vaughan, 1996; Dien et al., 2004). Sociological analysis of accident causation is gaining momentum as an effective

approach towards understanding the social and organisational causes of accidents (see, for example: Perrow, 1984; Vaughn, 1996; Hopkins, 2000).

Vaughn (1996) rejects the prevalent explanations (provided by traditional engineering techniques) of the cause of the *Challenger* accident and presents an alternative sociological explanation that explores much deeper root cause of the failure, and warns us of the risks involved in modern complex technological systems. The *Columbia* accident investigation report identifies a "broken safety culture" as a focal point of the accident's organisational causes (CAIB, 2003). Vaughan recognised similarities between the *Columbia* and *Challenger* accidents in that both accidents occurred due to organisational system failures, and presented a causal explanation that links the culture of production, the normalisation of deviance, and structural secrecy in NASA. (CAIB, 2003: Chap. 8).

This paper provides a review of key traditional accident modelling approaches and their limitations, and describes new system-theoretic approaches to the modelling and analysis of accidents in complex sociotechnical systems. An overview of traditional approaches, in particular event-based models, to accident modelling is given in Chapter 2, including its limitations to analyse accidents in modern complex systems. In Chapter 3, we discuss the nature and complexity of modern sociotechnical systems, describe Reason's organisational model of accident causation, and discuss the recent developments of systemic accident models. Two main systemic models, Rasumussen's risk management framework and AcciMap accident analysis technique, and Leveson's Systems Theoretic Accident Modelling and Processes approach are discussed in Chapters 4 and 5 respectively. The recent work on the application of formal methods, based on formal logics, to accident modelling and analysis is discussed in Chapter 6. In Chapter 7, we discuss the social, cultural and organisational factors in system accidents, and review sociological and organisational theories on safety and accident causation. Finally, we discuss future trends in the application and development of systemic accident models that consider the simultaneous interactions of technical, human, social, cultural and organisational aspects of modern complex systems.

2. Traditional Approaches to Accident Modelling

The historical development of accident models and various approaches to accident analysis and prevention have been discussed by engineers, scientists, psychologists and sociologists (see, for instance: Heinrich et al., 1980; Ferry, 1998; Johnson, 2003; Leveson, 1995; Hollnagel, 2004; Perrow, 1984; Vaughan 1996). Heinrich et al. (1980) and Ferry (1998) discuss accident causation models and techniques for accident analysis and prevention. Recently, Johnson (2003) has provided a comprehensive survey of traditional and modern accident modelling and analysis techniques. Leveson (2001) evaluates classic chain-of-events models using recent aerospace accidents. In the following sections, we give an overview of the chains-of-events models and their limitations in analysing modern complex technological systems.

2.1 Sequential Event-Based Models

Event-based models, also known as sequential accident models, explain accident causation as the result of a chain of discrete events that occur in a particular temporal order. One of the earliest sequential accident models is the Domino theory proposed by Heinrich (Heinrich et al., 1980). According to this theory there are five factors in the accident sequence: 1) social environment (those conditions which make us take or accept risks); 2) fault of the person; 3) unsafe acts or conditions (poor planning, unsafe equipment, hazardous environment); 4) accident; 5) injury. These five factors are arranged in a domino fashion such that the fall of the first domino results in the fall of the entire row (Figure 1). This illustrates that each factor leads to the next with the end result being the injury.

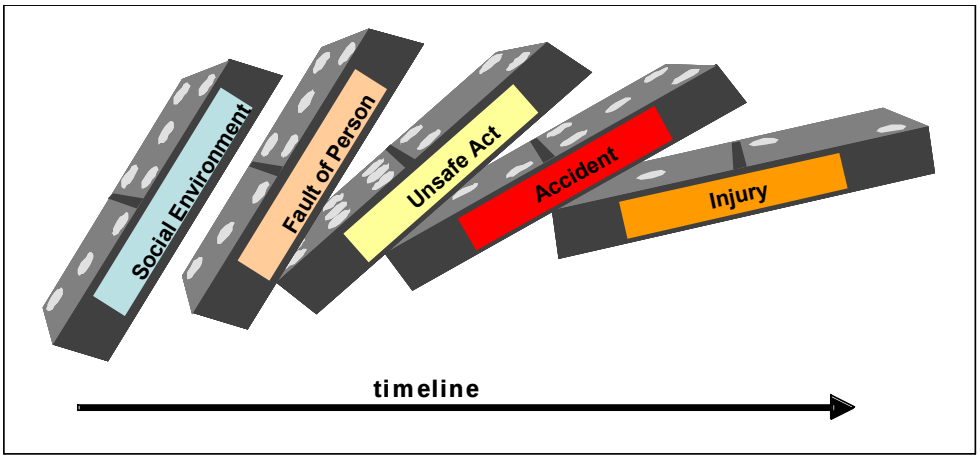

Figure 1: Heinrich's Domino Model of Accident Causation

An undesirable or expected event (the root cause) initiates a sequence of subsequent events leading to an accident. This implies that the accident is the result of a single cause, and if that single cause can be identified and removed the accident will not be repeated. The reality is that accidents always have more than one contributing factor.

2.2 Chains of Time-Ordered Events Models

Sequential models work well for losses caused by failures of physical components or human errors in relatively simple systems. While the Domino model considers only a single chain of events, event-based accident models can also be represented by multiple sequences of events in the form of hierarchies such as event tree and networks

A detailed description of these models can be found in Leveson (1995). The events considered in these models generally correspond to component failure, human error, or energy-related event. For example, in the Multilinear Events Sequencing (MES) model (Benner, 1975) every event is a single action by an actor. A timeline is included to show the timing sequencing of the events and conditions (Figure 2). Multiple chains of events, corresponding to different actors, are synchronised using the timeline. The MES charting method provides criteria to guide the development of the explanation of specific accidents in a manner that facilitates the transfer of knowledge among accident investigators. The accident sequence begins when a stable situation is disturbed. If the actor involved in the sequence adapts to the disturbance, the accident is averted. Countermeasures can be formulated by examination of each individual event to see where changes can be introduced to alter the process.

Although the MES model shows how events are related and combine to cause accidents, the development and analysis of such models is time consuming and requires significant analyst expertise. Insensitivity of the analyst to the possibility of missing information has been shown to cause overconfidence in model predictions (Fischoff et al., 1978).

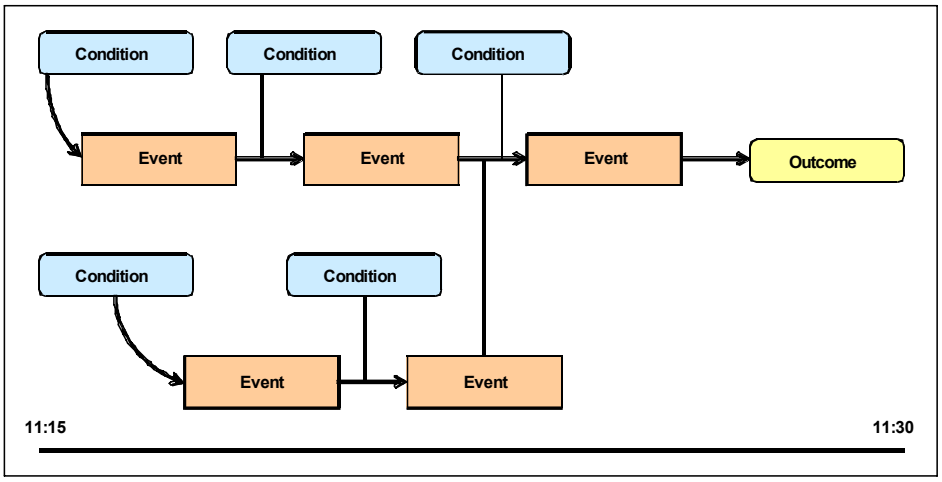

Figure 2: Activity events and outcomes for two actors including events (Ferry, 1988)

In event-based models, the events have a direct linear relationship. These models can only describe linear causality, and it is difficult to incorporate non-linear relationships. The first event in the chain is often considered the "initiating event"; however, the selection of the initiating event is arbitrary and previous events and conditions could always be added (Leveson, 2001). A particular event may be selected as the cause because it is the event immediately preceding the accident. The friendly fire shoot down of the two US Black Hawk helicopters in Iraq (AAIB, 1994) could be blamed on the F-15 pilots as the root cause, since the last condition before the accident was the firing of the missiles. However, the accident report has shown that there were a large number of factors and events that contributed to the accident. One reason for this tendency to look for a single cause is to assign blame, often for legal purposes. Occasionally, an accident investigator will stop at a particular event or condition that is familiar and can be used as an acceptable explanation of the accident. Usually there is no objective criterion for distinguishing one factor or several factors from the other factors that make up the cause of the accident (Leveson, 2001).

2.3 Risk Analysis Models

In many systems engineering areas, complex and safety critical systems development employ hazard analysis techniques to predict the occurrence of accidents in order to reduce risk and ensure safety in system design and operation. Hazard analysis is an activity by which sequences of events that can lead to hazards or accidents are identified, and the chance of such a sequence occurring is estimated (Leveson, 1986; ATEA, 1998). Leveson evaluates a number of models and techniques that are used in accident investigations and occasionally in predictive analysis. We discuss two widely used hazard analysis techniques that are employed during the early stages of system design.

Fault Tree Analysis

Fault Tree Analysis (FTA) is primarily a technique for analysing the causes of hazards, and traditionally used for the safety analysis of electromechnical systems. A fault tree is a logical diagram that shows the relationship between a system failure, i.e. a specific undesirable hazardous event in the system, and failures of the components of the system. The component failures can be events associated with hardware, software and human error. It is a technique based on deductive logic. The analyst first assumes a particular system state, and a top (hazardous) event and then identifies the causal events (component failure) related to the top event and the logical relations between them, using logic symbols to describe the relations. A fault tree is a simplified representation of a very complex process. It does not convey any notion of time ordering or time delay. A fault tree is a snapshot of the state of the system at one point in time.

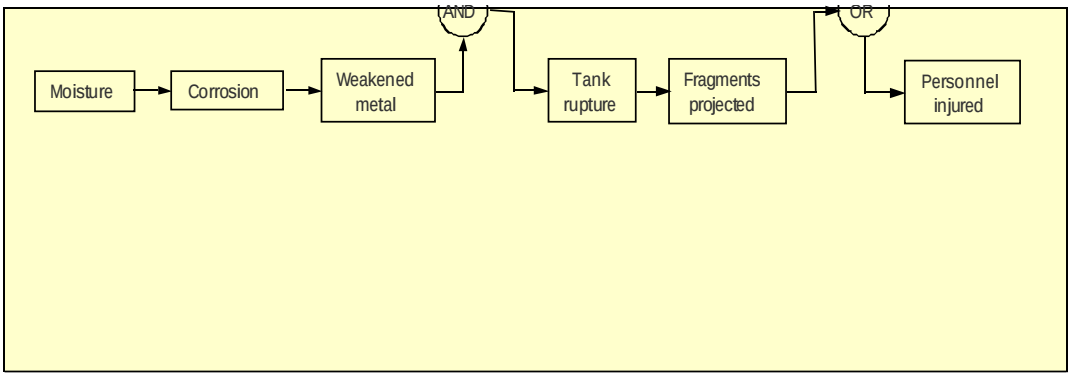

Figure 3: Fault-Error-Failure Model (Leveson, 1995)

Failure Modes and Effects Analysis

Failure Modes and Effects Analysis (FMEA) was originally developed to predict the reliability of hardware systems. The objective of the analysis is to validate the design by listing all possible sources of failures of a system's components and by determining the various effects of these failures on the behaviour of the system. FMEA uses forward search based on an underlying chain-of-events model, where the initiating events are failures of individual components. FMEA is most appropriate for standard components with few and well-known failure modes, and is effective for analysing single point failure modes. FMEA considers each failure as an independent occurrence with no relation to other failures in the system. Thus this technique does not consider multiple or common cause failures, and it is quite difficult to investigate accidents that could

arise due to combination of failure modes. It cannot easily be used to analyse the interactions between complex subsystems. Furthermore, the analysis is static, i.e., real-time aspects are ignored. Because FMEAs establish the end effects of failures, they are sometimes used in safety analysis for predicting the failures and hazards that may lead to accidents. Failure Modes and Effects Criticality Analysis (FMECA) is basically an FMEA with more detailed analysis of the criticality of the failure.

FTA, FMEA, and FMECA are standard risk analysis methods for component failure analysis. Such traditional approaches have serious limitations in the analysis of complex sociotechnical systems, since they do not consider the organisational, social, and complex interactions between the various system components.

Sequential models assume that the cause-effect relation between consecutive events is linear and deterministic. Analysing an accident may show that cause *A* led to effect *B* in a specific situation, while *A* may be a composite event (or state) in turn having numerous causes (Hollnagel, 2001). Thus, these models cannot comprehensively explain accident causation in modern sociotechnical systems where multiple factors combine in complex ways leading to system failures and accidents.

3. Modern Approaches to Accident Modelling

3.1 Complexity of Sociotechnical Systems

In modern complex systems, humans interact with technology and deliver outcomes as a result of their collaboration; such outcomes cannot be attained by either the humans or technology functioning in isolation. Such systems, composed of human agents and technical artefacts, are often embedded within complex social structures such as the organisational goals, policies and culture, economic, legal, political and environmental elements. Sociotechnical theory implies that human agents and social institutions are integral parts of the technical systems, and that the attainment of organisational objectives are not met by the optimisation of the technical system, but by the joint optimisation of the technical and social aspects (Trist & Bamforth, 1951). Thus, the study of modern complex systems requires an understanding of the interactions and interrelationships between the technical, human, social and organisational aspects of the system.

For example, civil aviation is a complex public transportation system comprising technological artefacts (aircrafts, runways, luggage transport systems, communication equipment, etc.); these artefacts have various interconnections and relationships and they all play an essential role in the functioning of this transport system as a whole (Kroes et al., 2006). These technical artefacts and systems operate in a social-organisational environment which constitutes various policies and procedures, the air traffic control system, legal and economic aspects. Thus, the functioning of this transport system is also dependent on the functioning of social elements and on the behaviour of various human agents, and not purely on the functioning of the technical artefacts.

Charles Perrow's seminal work on normal accident theory (Perrow, 1984) provides an approach to understanding accident causation in complex organisations managing hazardous technologies such as nuclear power plants, petrochemical plants, aircraft, marine vessels, space, and nuclear weapons. Perrow analyses many notable accidents involving complex systems such as the 1979 Three Mile Island nuclear power accident, and identifies that the characteristics that make a technological system or organisations more prone to accident are complex interactions and tight coupling.

A complex system is composed of many components that interact with each other in linear and complex manners. Linear interactions are those that are expected in production or maintenance sequences, and those that are quite visible even if unplanned (during design), while complex (nonlinear) interactions are those of unfamiliar sequences, unplanned and unexpected sequences, and either not visible or not immediately comprehensible (Perrow, 1984). Two or more discrete failures can interact in unexpected ways which designers could not predict and operators cannot comprehend or control without exhaustive modelling or test.

The type of coupling (tight or loose coupling) of components in a system affects its ability to recover from discrete failures before they lead to an accident or disaster. Perrow (1984) discusses the characteristics of tightly and loosely coupled systems. Tightly coupled systems have more time-dependant processes, so that is a failure or event in one component has an immediate impact on the interacting component. Tightly coupled systems have little slack, quantities must be precise and resources

cannot be substituted for one another. For example, a production system must be shutdown if a subsystem fails because the temporary substitution of other equipment is not possible. In contrast, loosely coupled systems are more forgiving; delays are possible, products can be produced in a number of ways, and slack in resources is possible.

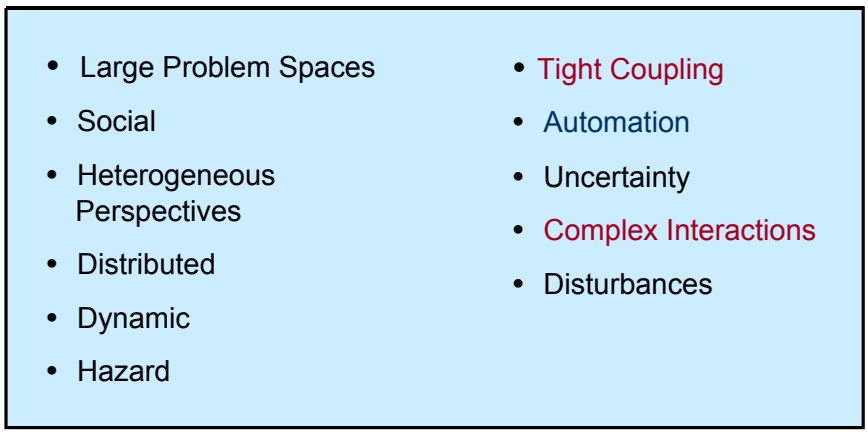

Figure 4: Complexity in Sociotechnical Systems: A Multidimensional Concept (Perrow, 1984; Vicente, 1999)

Vicente (1999) provides an excellent description of complexity in sociotechnical systems, and discusses a number of interrelated characteristics broadly present in different types of sociotechnical systems such as: large problem space, social interaction of groups of people, heterogeneous perspectives, distributed nature, dynamic properties, hazards in operations, automation, uncertainty in the data available to operators, and mediated interaction. Every sociotechnical system is different across its application domain and does not necessarily rate highly on all of these dimensions.

3.2 Reason's Organisational Model of System Accidents

Reason (1990; 1997) developed an organisational model for explaining accident causation in complex technological systems. Organisational accidents do not occur due to a single human error; rather they arise from the interconnection of several causal factors originating at many levels in an organisation. Reason emphasises the concept of organisational safety and how defences (protection barriers such as material, human and procedures) may fail. In this approach the immediate or proximal cause of the accident is a failure of people at the "sharp end" who are directly involved in the regulation of the process or in the interaction with the technology (Reason, 1990; Woods et al., 1994). Reason (1997) defines organisational accidents as situations in which latent conditions (arising from such aspects as management decision practices, or cultural influences) combine adversely with local triggering events (such as weather, location, etc.) and with active failures (errors and/or procedural violation) committed by individuals or teams at the sharp end of an organization, to produce the accident.

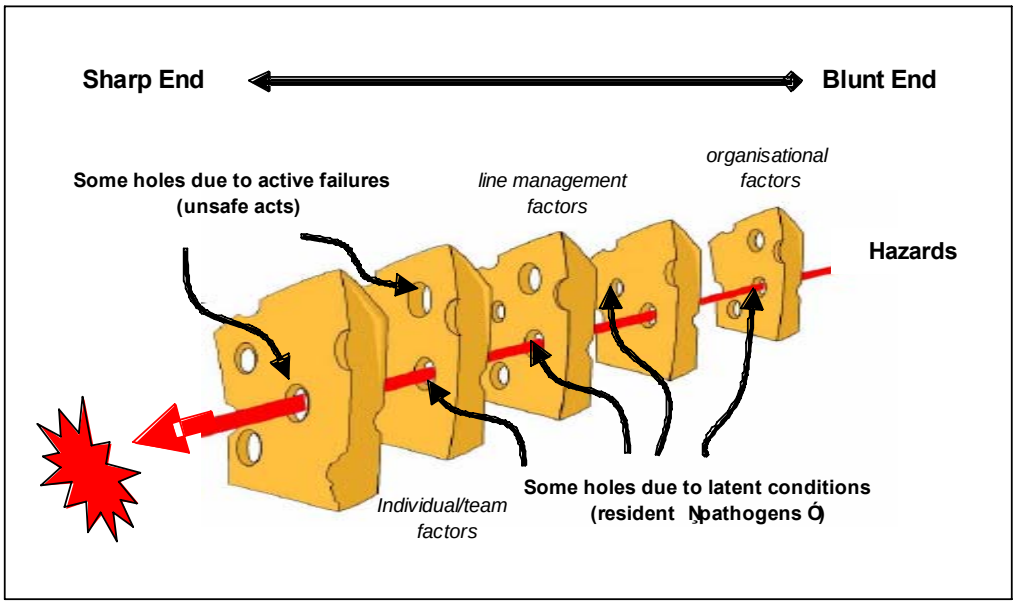

Figure 5: Swiss Cheese Model of Defences (Reason, 1997)

Defences, barriers, and safeguards occupy a key position in the system approach (Reason, 1997) to safety in complex systems. High technology systems have many defensive layers: some are engineered (alarms, physical barriers, automatic shutdowns, etc), others rely on humans (surgeons, anesthetists, pilots, control room operators, etc) behaviour, and yet others depend on procedures and administrative controls. Reason's model is based on the defences in depth philosophy from military and nuclear power plant industry, that is a defensive system that involves many layers of barriers, each designed to support the other in order to reduce the likelihood of occurrence of an accident or disaster. The dynamics of accident causation are represented in the Swiss cheese model of defences (Figure 5), which shows an accident emerging due to holes in barriers and safeguards.

In an ideal world all defensive layers should be intact allowing no penetration to happen. However, in the real world: defences may deteriorate over time, such as the corroded sprinklers on the Piper Alpha accident; modification or redesign may weaken or eliminate defences; defences can be removed during calibration, maintenance and testing, or as a result of errors and violations (Reason, 1997). The control room operators of the Chernobyl nuclear reactor successively removed layers of defence in order to complete their task of testing a new voltage generator. In reality, however, they are more like slices of Swiss cheese, having many holes; though unlike in the cheese, these holes are continually opening, shutting, and shifting their location. The presence of holes in any one "slice" does not normally cause a bad outcome. Usually, this can happen only when the holes in many layers momentarily line up to permit a trajectory of accident opportunity, bringing hazards into damaging contact with victims (Figure 5.). The holes in the defences arise for two reasons: active failures and latent conditions. Nearly all adverse events involve a combination of these two sets of factors.

Active failures are the unsafe acts committed by people who are in direct contact with the patient or system. They take a variety of forms: slips, lapses, fumbles, mistakes, and procedural violations (Reason, 1990). Active failures have a direct and usually short lived impact on the integrity of the defences. At Chernobyl, for example, the operators wrongly violated plant procedures and switched off successive safety systems, thus

creating the immediate trigger for the catastrophic explosion in the core. Followers of the person approach often look no further for the causes of an adverse event once they have identified these proximal unsafe acts. But, as discussed below, virtually all such acts have a causal history that extends back in time and up through the levels of the system.

Latent conditions are the inevitable "resident pathogens" within the system (Reason, 1997). They arise from decisions made by designers, builders, procedure writers, and top-level management. Such decisions may be mistaken, but they need not be. All such strategic decisions have the potential for introducing pathogens into the system. Latent conditions have two kinds of adverse effect: they can translate into error provoking conditions within the local workplace (for example, time pressure, understaffing, inadequate equipment, fatigue, and inexperience) and they can create long-lasting holes or weaknesses in the defences (untrustworthy alarms and indicators, unworkable procedures, design and construction deficiencies, etc). Latent conditions, as the term suggests, may lie dormant within the system for many years before they combine with active failures and local triggers to create an accident opportunity. Unlike active failures, whose specific forms are often hard to foresee, latent conditions can be identified and remedied before an adverse event occurs. Understanding this leads to proactive rather than reactive risk management.

The notion of latent factors supports the understanding of accident causation beyond the proximate causes, which is particularly advantageous in the analysis of complex systems that may present multiple-failure situations. This model has been particularly useful in accident investigation, as it addresses the identification of latent failures within the causal sequence of events as well. This model has been widely applied in many domains to understand how accidents are caused such as the oil and gas industry (Wagenaar et al., 1994), commercial aviation (Maurino et al., 1995), and it has become a standard in medicine (Reason et al., 2000; Reason, 2000).

This model places a great emphasis on the search for latent or organisational causes and provides an understanding of how these are related to the immediate causes at the sharp end. Reason (1990) conducted a number of case studies of the Three Mile Island, Bhopal disaster, and Chernobyl accident and identified several latent failures related to organisational, management and design failures. In the Swiss cheese model the latent and active errors are causally linked to management as a linear sequence of events, and this can lead to the illusion that the roots of all accidents or even errors stem from the organisation's management. Shorrock et al. (2003) argue that, in some cases, the main contributory factors might well have been active errors with more direct implications for the accident causation.

Johnson & Botting (1999) employed Reason's model to understand the organisational aspects of the Watford Junction railway accident. They studied the latent conditions that contributed to the active failure by the train driver to violate two sets of signals. Numerous organisational factors were identified as the causal factors that contributed to the probability of the accident (see Table 1). However, this model does not give a clear explanation how these causal factors combined to provide the circumstances for an accident to take place. For example, the main defences of the Watford Junction, the positioning of the Permanent Speed Restrictions signs and the junction signals, were not independent, and Johnson & Botting recommend the use of formal methods to analyse this complexity in detail. Furthermore, the causal links between distant latent

conditions (organisational factors) and the accident outcome is complex and loosely coupled (Shorrock et al., 2003), and Reason's model only guides to a high-level analysis of the contributory factors involved in an organisational accident.

Table 1: Watford Junction Railway Accident – Active failures and latent conditions (Johnson & Botting, 1999)

Active failures	Latent Conditions
1. Driver violates signals and horns	1. Speed boards placed in wrong positions 2. Drivers never given information on why speed boards are there 3. There had been several Signals Passed at Danger (SPADs) but a signal sighting committee was not convened 4. An inspection of the signalling system at Watford Junction was never carried out 5. The driver did not know of the reduced overlap between signals 6. The driver had committed SPADs recently and so should have been in the 'at risk' category

However, epidemiological models still follow the principles of sequential models (Hollnagel, 2004) as they show the direction of causality in a linear fashion. The theory behind the Swiss cheese model does not define in sufficient detail what the system failures or holes in the cheese are, and Shappell & Weigmann (2000) have developed a framework based on Reason's model which assists investigators to examine human error in the field and to track those factors (the holes in the cheese) responsible for the accidents as well. Shappell & Weigmann's framework is not based on any particular theoretical framework but based on aviation accidents, and may not be easily transferable to other domains.

Reason's model shows a static view of the organisation; whereas the defects are often transient i.e. the holes in the Swiss cheese are continuously moving. In reality, the sociotechnical system is more dynamic than the model suggests.

3.3 Integrating Event-Chain and Reason's Models

The event chain (fault-error-failure) model was originally designed with the objective of describing the propagation of faults in technical systems. Conversely, Reason's (1990; 1997) Swiss cheese model was intended to describe the organisational factors and their causal relationships to front end operator errors leading to an accident.

In sociotechnical systems, computers and technical artifacts in general are being more and more tightly integrated with human activities. Failures in sociotechnical systems

are the result of a combination of factors meshed into a complex causal network spread over several hierarchical levels within an organisation (Reason, 1990; 1997). Besnard & Baxter (2003) argue that technical and organisational issues need to be simultaneously considered to capture the causal mesh leading to accidents and discuss the integrative representation of the event chain and Reason's Swiss cheese models. There are strong common ideas between these two models (Besnard & Baxter, 2003):

- *Systems can be decomposed into layers*. Each layer represents a sub-system, a state or an actor that has an impact on the functioning of the entire system.
- *Failures wait for calling conditions*. Some unstable conditions can be present in a given system without having any immediate effect. A failure, from this point of view, is an unlikely combination of a number of contributing factors.
- *Events propagate*. Accidents are not caused by the occurrence of sudden unfavourable circumstances. Instead, they are generated by early design faults that, under certain conditions, trigger an undesired event.
- *Events escalate*. A combination of local failures accounts for the breakdown of full systems.

Besnard & Baxter (2003) state that each organisational layer invariably contains one or more holes, which can be attributed to the occurrence of fault-error-failure chains during its creation or functioning. One then gets an elementary failure generation chain for a hole in a given system's layer (Figure 6) that provides an identifiable causal path for each hole. In other words, this approach provides a mapping between failures and holes in the system's layers, where event chain and Reason's models can be turned into compatible representations of systems failures. This opens up a new area of application for the event chain model, that of sociotechnical system failures. Equally, it allows Reason's model to connect to technical causal paths of failures in systems.

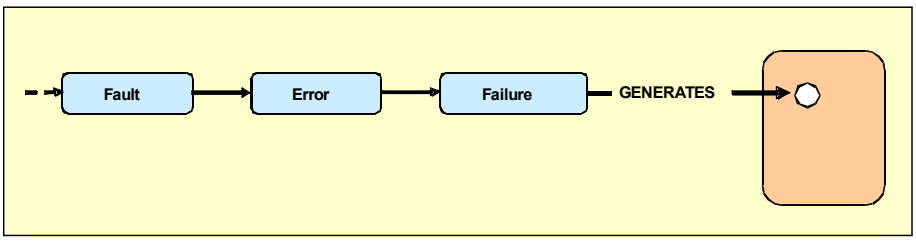

Figure 6: An elementary event chain generating a hole in a system layer (Besnard & Baxter, 2003)

The validity of this integrated model has been demonstrated by analysing the failure in the Therac-25 sociotechnical system. THERAC-25 was an X-ray treatment machine designed to destroy tumors in deep body tissues. Radiation overdoses happened between 1985 and 1987 and several patients died from subsequent injuries. The machine was recalled in 1987 for extensive design changes, including hardware safeguards against software errors (Leveson, 1993).

Besnard & Baxter (2003) developed a three-layer model for the THERAC-25 system: the regulation authorities, the company who developed the system, and the programmer who wrote the code, and introduced a fault-error-failure chain for each hole in the various system layers (see Figure 7). One of the many chains for each of the layers is described below:

- *The programmer* did not take all of the system's real-time requirements into account (fault).
- This led to the possibility of flaws in some software modules (error) that degraded the reliability of the software (failure).
- *The company* did not perform all the required tests on the software (fault). This resulted in bugs in some modules remaining undetected and hence unfixed (error), thereby triggering exceptions when the given modules were called (failure).
- *The regulation authorities* did not thoroughly inspect the system (fault). This led to some flaws remaining undetected (error). In turn, these flaws caused injuries and deaths when the system was used (failure).

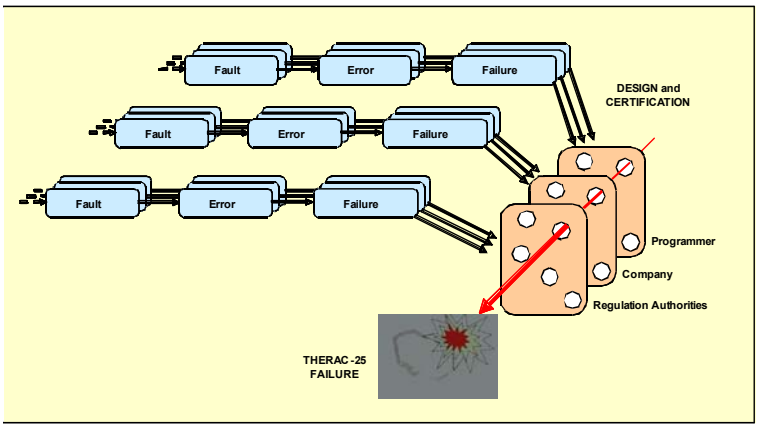

Figure 7: Integrating Event Chain and Reason's Models (Besnard & Baxter, 2003)

The resulting integrated model offers a richer description of sociotechnical failures by suggesting a mapping between sequences of events (a fault-error-failure chain) and holes in the layers of a system (Reason's Swiss cheese model). This approach provides some intrinsic interest since it constitutes a step forward in reconciling technical and organisational views on failures in sociotechnical systems.

3.4 Systemic Accident Models

3.4.1 Systems Theoretic Approach

New approaches to accident modelling adopt a systemic view which considers the performance of the system as a whole. In systemic models, an accident occurs when several causal factors (such as human, technical and environmental) exist coincidentally in a specific time and space (Hollnagel, 2004). Systemic models view accidents as emergent phenomena, which arises due to the complex interactions between system components that may lead to degradation of system performance, or result in an accident.

Systemic models have their roots in systems theory. Systems theory includes the principles, models, and laws necessary to understand complex interrelationships and interdependencies between components (technical, human, organisational and management) of a complex system.

In a systems theory approach to modelling, systems are considered as comprising interacting components which maintain equilibrium through feedback loops of

information and control. A system is not regarded as a static design, but as a dynamic process that is continually adapting to achieve its objectives and react to changes in itself and its environment. The system design should enforce constraints on its behaviour for safe operation, and must adapt to dynamic changes to maintain safety. Accidents are treated as the result of flawed processes involving interactions among people, social and organisational structures, engineering activities, and physical and software system components (Leveson, 2004).

Rasmussen adopts a system oriented approach based on a hierarchical sociotechnical framework for the modelling of the contextual factors involved in organisational, management and operational structures that create the preconditions for accidents (Rasmussen, 1997; Rasmussen & Svedung, 2000). Leveson (2004) proposes a model of accident causation called STAMP (Systems-Theoretic Accident Model and Processes) that considers the technical, human and organisational factors in complex sociotechnical systems.

3.4.2 Cognitive Systems Engineering Approach

Modern technology has changed the nature of human work from mainly manual tasks to predominantly knowledge intensive activities and cognitive tasks. Technology-driven approaches to automation have created new problems for human operator performance and new kinds of failure modes in the overall human-machine systems, which have led to many catastrophic accidents in the fields of aviation, nuclear power plants and military command and control (Parasuraman, 1997). This has influenced the development of new approaches for human performance and error modelling, and accident analysis of joint human-machine systems.

Cognitive systems engineering (Hollnagel & Woods, 1983) has emerged as a framework to model the behaviour of human-machine systems in the context of the environment in which work takes place. The traditional view is that "human errors" represent a *post hoc* rationalization (Woods et. al., 1994), which is based on the inverse causality principle: "if there is an effect, then there must be a cause". Cognitive systems engineering instead suggests that we cannot understand what happens when things go wrong without understanding what happens when things go right (Hollnagel & Woods, 2005). Hollnagel & Woods introduce a new paradigm on Joint Cognitive Systems which describes how humans and technology function as joint systems, rather than how humans interact with machines. Efforts to make work safe should start from an understanding of the normal variability of human and Joint Cognitive Systems performance, rather than assumptions about particular, but highly speculative "error mechanisms" (for a detailed discussion see: Hollnagel & Woods, 2005).

Two systemic accident models for safety and accident analysis have been developed based on the principles of cognitive systems engineering: the Cognitive Reliability and Error Analysis Method (CREAM); and the Functional Resonance Accident Model (FRAM).

CREAM is based on the modelling of cognitive aspects of human performance for an assessment of the consequences of human error on the safety of a system (Hollnagel, 1998). Two versions of CREAM have been developed for accident modelling: DREAM (Driver Reliability and Error Analysis Method) for analysis of traffic accidents; and BREAM for use in maritime accident analysis (Hollnagel, 2006a).

FRAM is a qualitative accident model that describes how functions of system components may resonate and create hazards that can run out of control and lead to an accident (Hollnagel, 2004). FRAM is based on the premise that performance variability, internal variability and external variability are normal, in the sense that performance is never stable in a complex sociotechnical system such as aviation.

4. Rasmussen's Sociotechnical Framework for Risk Management

The complexity and rapid advancements in technology have led to the development of high-risk sociotechnical systems, which are managed by complex organisations operating in highly volatile and dynamic environmental conditions such as market competition, economic and political pressures, legislation and increasing social awareness of safety (Rasmussen, 1997). Rasmussen postulates that these factors have transformed the dynamic character of modern society and continuously influence the work practices and human behaviour in the operation of complex systems. Deterministic (e.g. sequential chain-of-events) causal models are inadequate to study failures and accidents in highly adaptable sociotechnical systems. Rasmussen adopts a system oriented approach and proposes a framework for modelling the organisational, management and operational structures that create the preconditions for accidents. In this section, we describe Rasmussen's conceptual control framework for modelling risk management in complex sociotechnical systems.

4.1 Structural Hierarchy and System Dynamics

Rasmussen's framework for risk management has two parts: Structure and Dynamics.

Structural Hierarchy

Rasmussen views risk management as a control problem in the sociotechnical system, where human injuries, contamination of environment, and loss of investment occur due to loss of control of physical processes. The activity of people in their work environment can trigger an accidental flow of events or change the normal operational flow that can result in an accident. Safety, then, depends on the control of work processes so as to avoid accidental side effects causing harm to people, environment, or investment (Rasmussen, 1997).

The sociotechnical system involved in risk management includes several hierarchical levels ranging from legislators, organisation and operation management, to system operators. Figure 8 provides a representative example, although the precise number of levels and their labels can vary across industries.

The top level L1 describes the activities of government, who through legislation control the practices of safety in society. Level L2 describes the activities of regulators, industrial associations and unions (such as medical and engineering councils) that are responsible for implementing the legislation in their respective sectors. Understanding these two levels usually requires knowledge of political science, law, economics and sociology. Level L3 describes the activities of a particular company, and usually requires knowledge of economics, organisational behaviour, decision theory and sociology. Level L4 describes the activities of the management in a particular company that lead, manage and control the work of their staff. Knowledge of management theories and industrial-organisational psychology is used to understand this level. Level L5 describes the activities of the individual staff members that are interacting directly with technology or process being controlled such as power plant control operators, pilots, doctors and nurses. This level requires knowledge in new disciplines such as psychology, human-machine interaction and human factors. The bottom level

L6 describes the application of engineering disciplines involved in the design of potentially hazardous equipment and operating procedures for process control such as nuclear power plant and aviation. Understanding this level usually requires knowledge of science and various engineering disciplines.

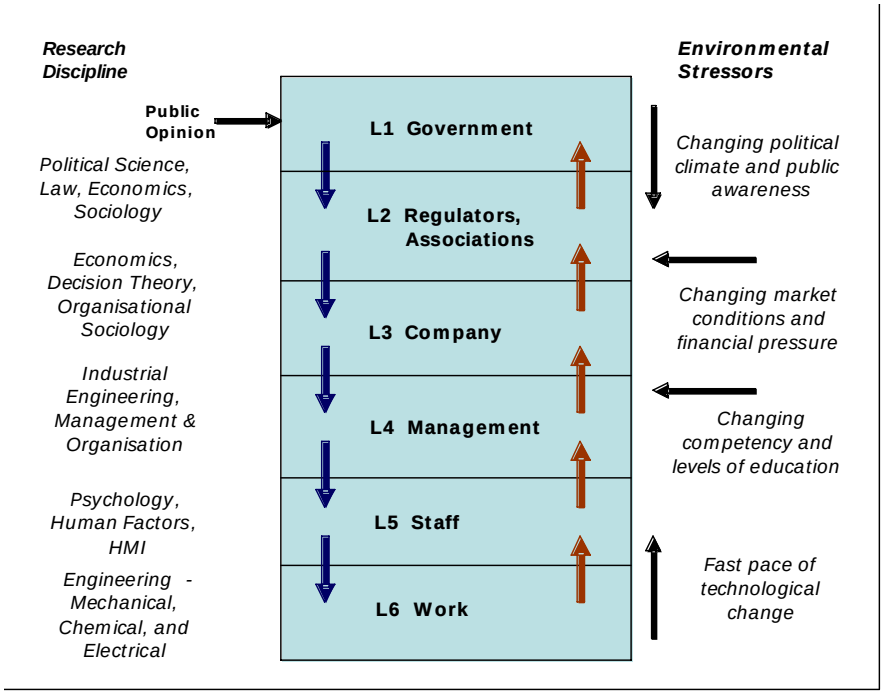

Figure 8: Hierarchical Model of Sociotechnical System involved in Risk Management (Rasmussen, 1997)

For example, in the context of health care, this hierarchy would include, from bottom to top: patients, providers (doctors and nurses), department managers, hospital CEOs, professional regulators and associations, government (i.e., civil servants and politicians), and the media (CEL, 2007). Each of these individuals and stakeholders makes decisions that affect patient safety.

Traditionally, each level is studied separately by a particular academic discipline, for example, risk management at the upper levels is studied without any detailed consideration of processes at the lower levels. This framework points to a critical factor that is overlooked by all horizontal research efforts, that is, the additional need for "vertical" alignment across the levels in Figure 8. The organisational and management decisions made at higher levels should transmit down the hierarchy, whereas information about processes at lower levels should propagate up the hierarchy. This vertical flow of information forms a closed loop feedback system, which plays an essential role in the safety of the overall sociotechnical system. Accidents are caused by decisions and actions by decision makers at all levels, and not just by the workers at the process control level.

As shown on the right of Figure 8, the various layers of complex sociotechnical systems are increasingly subjected to external disruptive forces, which are unpredictable, rapidly changing and have a powerful influence on the behaviour of the sociotechnical system. When different levels of the system are being subjected to different pressures, each operating at different time scales, it is imperative that efforts to improve safety within a level be coordinated with the changing constraints imposed by other levels.

System Dynamics

In complex dynamic environments it is not possible to establish procedures for every possible condition, in particular for emergency, high risk, and unanticipated situations (Rasmussen, 1997). In nuclear power plants, where tasks and procedures are strictly prescribed, violations of instructions have been repeatedly observed (Vicente et al., 2001; Vicente et al., 2004). Vicente argues that operator's violation of formal procedures appear to be quite rational (sensible) given the actual workload and timing constraints. The behaviour of operators is context dependent and is shaped by the dynamic conditions in the work environment.

Decision making and human activities are required to remain between the bounds of the workspace defined by administrative, functional and safety constraints. Rasmussen argues that in order to analyse a work domain's safety, it is important to identify the boundaries of safe operations and the dynamic forces that may cause the sociotechnical system to migrate towards or cross these boundaries. Figure 9 shows the dynamic forces that can cause a complex sociotechnical system to modify its structure and behaviour over time.

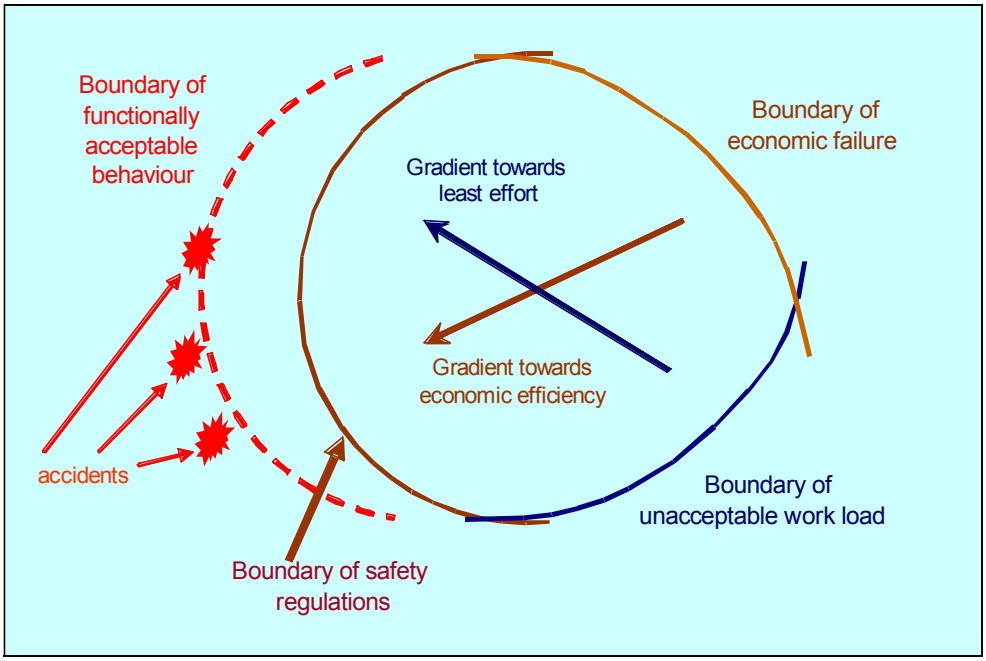

Figure 9: Boundaries of Safe Operation (Rasmussen, 1997)

The safe space of performance within which actors can navigate freely is contained within three boundaries: individual unacceptable workload; financial and economic constraints; and the safety regulations and procedures. The financial pressures produce a cost gradient that influences individual human behaviour to adopt more economically effective work strategies; while workload pressures result in an effort gradient motivating individuals to change their work practices to reduce cognitive or physical work. These gradients induce variations in human behaviour that are analogous to the "Brownian movements" of the molecule of a gas. The financial and psychological pressures cause people to change the way in which they perform their job, and may also lead to more adaptive and innovative ways of getting the task done.

Over a period of time, this adaptive behaviour causes people to cross the boundary of safe work regulations and leads to a systematic migration toward the boundary of functionally acceptable behaviour. This may lead to an accident if control is lost at the boundary. The migration in work practices does not usually have any visible, immediate threat to safety prior to an accident, because violation of procedures does not immediately lead to a catastrophe. At each level in the sociotechnical hierarchy, people are working hard, striving to respond to cost-effective pressures, but they do not see how their decisions interact with those made by other actors at different levels in the system (Woo & Vicente, 2003). Rasmussen asserts that these uncoordinated attempts of adapting to environmental stressors are slowly but surely "preparing the stage for an accident".

The safety control structure often changes over time, which accounts for the observation that accidents in complex systems frequently involve a migration of the system towards a state where a small deviation (in the physical system or operator behaviour) can lead to a catastrophe. The analyses of several accidents such as Bhopal and Chernobyl demonstrate that they have not been caused by coincidence of independent failures and human errors, but by a systematic migration of organisational behaviour towards an accident under the influence of pressure toward cost-effectiveness in an aggressive, competitive environment (Rasmussen, 1997).

Rasmussen's approach for improving safety and risk management raises the need for the identification of the boundaries of safe operation, making these boundaries visible to the actors and giving opportunities to control behaviour at the boundaries.

4.2 AcciMap Accident Analysis Technique

The AcciMap accident analysis technique is based on Rasmussen's risk management framework (described in the previous section). In the AcciMap technique, models in terms of functional abstraction are developed as they are more suitable for capturing the dynamics of highly adaptable sociotechnical systems (Rasmussen & Svedung, 2000). These models describe the information flow structure within the entire system involved in the control of hazardous processes. Rasmussen & Svedung recommend a sequence of phases of accident analysis, based on the risk management framework, together with a set of graphic representations useful to structure the analyses of hazardous work systems:

1) Selection and Analysis of Accident Cases
2) Identification of Actors
3) Development of a Generic AcciMap
4) Work Analysis

Selection and Analysis of Accident Cases

A representative set of accident cases are selected for the industrial sector under investigation. For each of these accident scenarios the causal chains of events are then analysed. From here an overview of the patterns of accidents related to a particular activity or system is generated by a cause-consequence analysis that is represented by a cause-consequence chart.

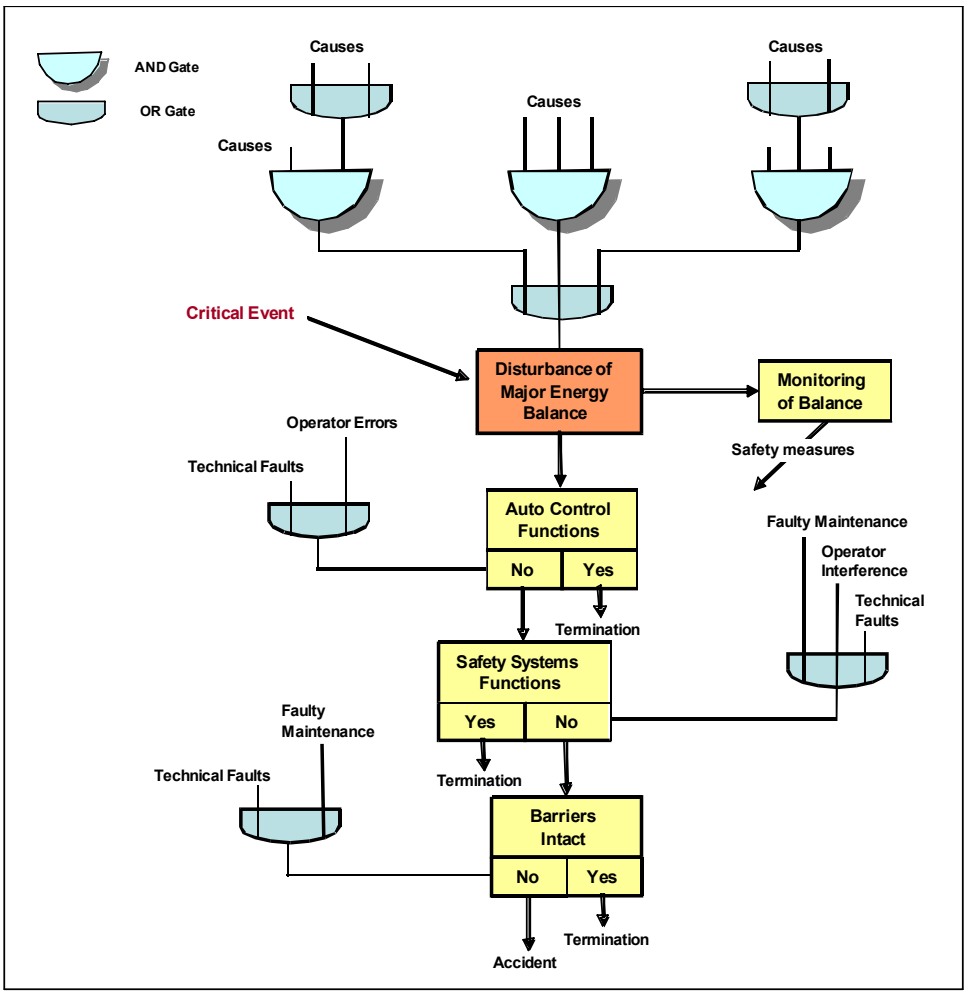

Figure 10: Cause-consequence diagram with multiple safety barriers (Rasmussen & Svedung, 2000)

A cause-consequence chart represents a generalisation that aggregates a set of accidental courses of events. Cause-consequence charts have been widely used as a basis for predictive risk analysis (Leveson, 1995). The set of events in a cause-consequence chart is determined by the choice of the critical event, which reflects the release of a well-defined hazard source, such as ”loss of containment of hazardous substance”, or ”loss of control of accumulated energy”. The *critical event* connects the causal tree (the logic relation among potential causes) with a consequent event tree (the possible functional and temporal relation among events) explicitly reflecting the switching of the flow resulting from human decisions or by automatic safety systems. (Svedung & Rasmussen, 2002). Figure 10 depicts a cause-consequence diagram which represents the anatomy of accidents in an industrial process plant with multiple safety barriers.

Identification of Actors

The cause-consequence chart focuses on the control of the hazardous process at the lowest level of the sociotechnical system (level 6 in Figure 8). In order to conduct a vertical analysis across the hierarchical levels, the cause-consequence chart representation is extended which explicitly includes the normal work decisions at the

higher levels of the sociotechnical system (levels 1-6 in Figure 8). This extension results in an *AcciMap* which shows the activities of various decision makers contributing to or preventing an accident. The AcciMap represents a mapping of these contributing factors onto the respective levels of a complex sociotechnical system identified in Figure 8.

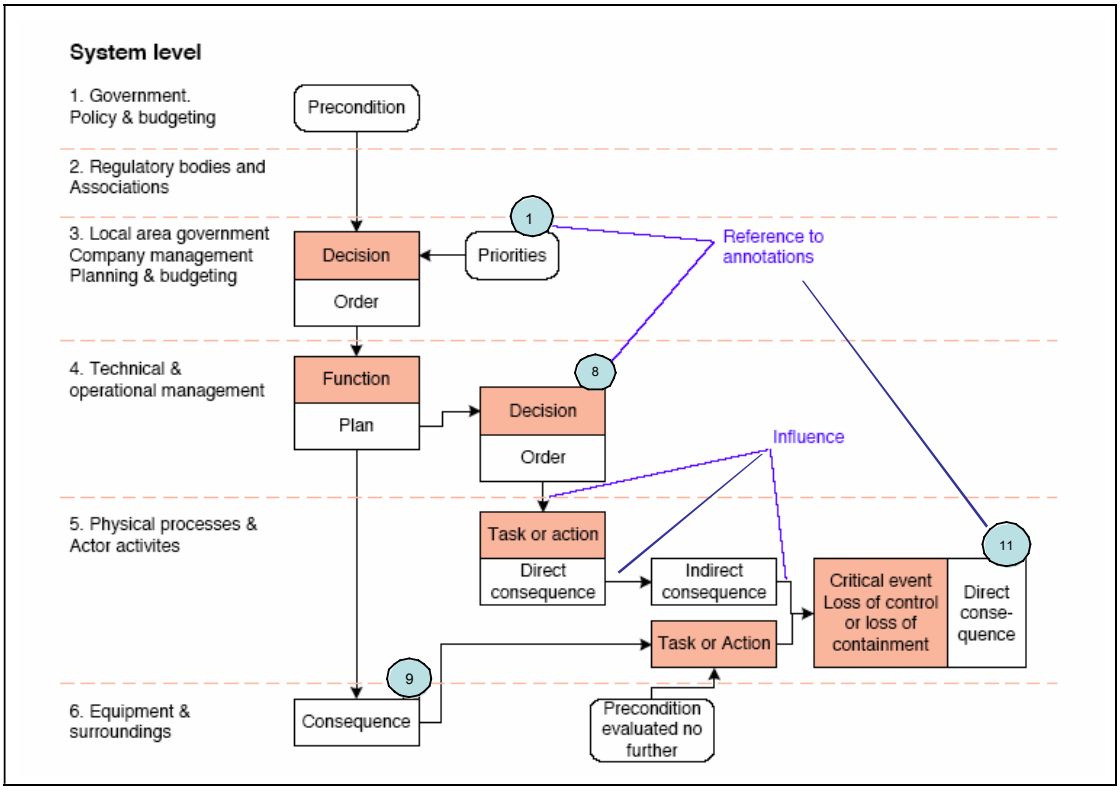

Figure 11: AcciMap Structure and Symbols (Rasmussen & Svedung, 2000)

The basic AcciMap is developed from analysis of one particular accident case, i.e., it reflects one particular course of events. The layout and symbols used in an AcciMap are shown in Figure 11 (Rasmussen & Svedung, 2000):

- At the bottom is a level representing the topography of the accident scene: the configuration and physical characteristics of the landscape, buildings, equipment, tools, vehicles, etc. found at the location and involved in the accident.
- At the next higher level is represented the accident processes, that is, the causal and functional relations of the dynamic flow, described in terms of the cause-consequence charts convention. In the flow are included "Decision/Action" boxes connected to consequence boxes where the flow has been or could be changed by human (or automated) intervention.
- At the levels above this, the "Decision/Action" box symbol is used to represent all decision-makers that, through decisions in their normal work context, have influenced the accidental flow at the bottom.

In this way, the AcciMap serves to identify relevant decision-makers and the normal work situation in which they influence the occurrence of accidents. The focus is not on the traditional search for identifying the "guilty person", but on the identification of those people in the system that can make decisions resulting in improved risk

management and hence to the design of improved system safety. A number of AcciMaps reflecting the results of particular accident analysis based on official accident reports have been presented (see, for example: Rasmussen & Svedung, 2000; Svedung & Rasmussen, 2002).

Development of a Generic AcciMap

The basic AcciMap represents the flow of events from one particular accident. From the set of AcciMaps based on the set of accident scenarios, a generalised map, a *generic AcciMap* is developed that identifies the interaction among the different decision makers and the events leading to an accident. The generic AcciMap regarding the transport of dangerous goods is shown in (Rasmussen & Svedung, 2000; Svedung & Rasmussen, 2002), and an AcciMap of the F-111 Chemical Exposure accident is shown in the next section (see Figure 12).

Work Analysis

For each accident scenario, the decision-makers, planners, and actors who have been involved in the preparation of accidental conditions are identified and represented in an *ActorMap*. This map should identify the individuals and groups that are involved in an adverse event at all relevant levels of society shown in Figure 8. An ActorMap is an extract of the generic AcciMap showing the involved decision makers, and an ActorMap in the transport of dangerous goods case study is shown in Rasmussen & Svedung (2000).

An ActorMap gives an overview of the decision making bodies involved in the preparation of the ”landscape” through which an accidental flow of events may ultimately evolve. Based on this map, an *InfoMap* can be developed which represents the information flow among decision-makers during normal activities. This normative information system is used for a representation of the actual communication found within the particular workplace. The InfoMap can also identify weak links in the communication patterns with organisations, such as communication links that are not active or not adequately explicit. Rasmussen & Svedung (2000) have developed a number of InfoMaps and discuss their use for analysis of the communication among actors.

4.3 Causal Analysis of F-111 Chemical Exposure of RAAF Workers

For more than 20 years, since the late 1970s, Royal Australian Air Force (RAAF) maintenance personnel have been working inside the fuel tanks of F-111 aircraft, resealing leaking seams, in an ongoing series of repair programs. They worked in cramped and very unpleasant conditions, sometimes in unbearable heat and sometimes in near freezing temperatures, and they suffered chronic and occasionally acute exposure to the hazardous substances with which they worked. The resulting symptoms included skin rash, gastro-intestinal problems, headaches and loss of memory (Clarkson et al., 2001).

In early 2000, after the health of more than 400 maintenance workers had been seriously affected, RAAF finally realized the problem and the fuel tank repair program were suspended. This had a negative impact on the availability of F-111 aircraft, which resulted in a detriment to defence capability.

Initially, the material made available to the F-111 Board of Inquiry (BOI) points to ongoing failings at a managerial level to implement a safe system of work and co-ordinate processes within a complex organisation. The BOI hence pointed out that if anybody is to be held accountable, it should be the RAAF itself. The aim of the investigation, however, was not to assign blame; it was conducted to understand how the exposure occurred and to make recommendations designed to reduce the chance of recurrence.

A wide array of causal and contributory factors, occurring over 20 years, combined in complex ways to affect the health of hundreds of RAAF maintenance workers (Clarkson et al., 2001). A causal analysis was conducted for the spray seal program, and a causal diagram was developed based on Rasmussen's (1997) AcciMap technique. This analysis is based on the assumption that there is no ultimate cause or causes responsible for the accident; rather many causal factors contribute to the final outcome, including latent factors within the organisation as discussed in Reason's (1997) organisational accident model.

The causal diagram of the spray seal program (Figure 12) constitutes six hierarchical levels, where the principle employed is that the more remote the cause with respect to the final outcome, the higher up in the diagram it is located. The diagram is constructed by starting with the "Health outcomes" (the accident) and asking why it occurred, which leads to the identification of preceding causes. Counterfactual reasoning is employed to determine the necessary causal factor, in the sense that, had this factor been otherwise, the accident probably would not have occurred. The causal pathways are then determined, proceeding from the bottom of the diagram upwards (see Figure 12).

At the bottom are the outcomes - damage to the health of Air Force workers leading to suspension of the spray seal program and consequent reduction in the availability of F-111 aircraft. Next level up are the immediate causes. Above that are the organisational causes, to do with the way the Air Force as an organisation functioned. Above that are shown a number of Air Force values that accounted for many of the factors at the organisational level. Finally there are two levels, government and society, both beyond the Air Force organisation and over which the Air Force therefore has no control.

A summary of main findings and explanations of the various contributory factors and causal pathways is described in the BOI report (Clarkson et al., 2001: Chap. 11). Here, the causal paths leading to the failure of the chain of command to operate optimally is described.

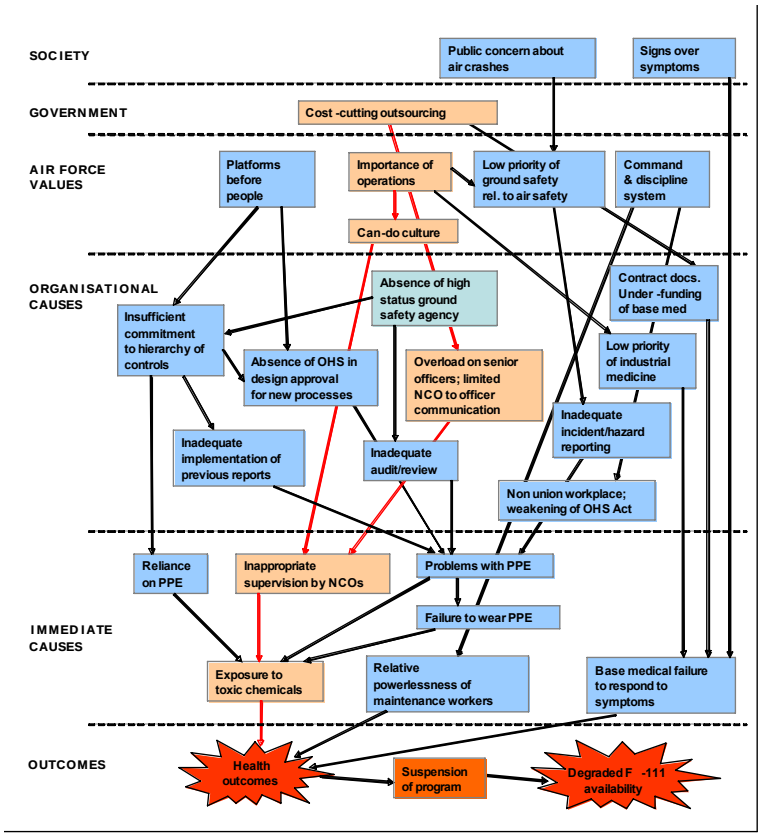

Figure 12: AcciMap of F-111 Seal/Reseal Program (Clarkson et al., 2001)

At the lowest level of the chain, non-commissioned officers put up with a variety of inadequacies in Personal Protective Equipment (PPE) as well as other equipment failures and ventilation problems, without raising these matters effectively through the chain of command, or in some other way. They did so in part because of the pressure that they perceived operational requirements placed on them to get the job done as quickly as possible, which resulted in a well-intentioned but inappropriate can-do response. They also often failed to take matters higher up because of the expectation that whenever possible they should resolve things at their level.

There was a particular weak link in the chain of command between the senior NCOs and the junior engineering officers, and there was limited communication between these two levels. Part of the reason for this was the very broad span of responsibilities which junior engineering officers were expected to shoulder. This in turn was a consequence of reductions in staff numbers as part of a general downsizing. Senior officers, too, were suffering extreme work overload as a result of being expected to carry out market testing (outsourcing) functions as well as their normal supervisory functions. The result was that senior officers had relatively little conception of what was occurring on the hanger floor. These weaknesses at the upper levels of the chain of command stem fairly directly from government policy decisions lying largely outside the control of the Air Force.

The causal diagram in Figure 12 is based on the official F-111 Board of Inquiry report (Clarkson et al., 2001). The causal flow diagram looks at the culture of RAAF as well as factors that lie beyond the organisational limits of RAAF. This analysis concludes that the failure of the chain of command to operate optimally predominantly lies at the values and culture of RAAF, and to government policies such as the government

initiated cost-cutting and down-sizing of employees, and social attitudes such as the focus on air safety driven partly by public pressure.

In this way, the causal analysis serves to identify relevant decision-makers and the normal work situation in which they influence and condition possible accidents. The focus is not on the traditional search for identifying the "guilty person", but on the identification of those people in the system that can make decisions resulting in improved risk management and hence to the design of improved system safety.

5. Systems Theoretic Accident Model and Processes (STAMP)

5.1 Basic Concepts in STAMP

Leveson (2004) proposes a model of accident causation that considers the technical (including hardware and software), human and organisational factors in complex sociotechnical systems. According to Leveson, "The hypothesis underlying the new model, called STAMP (Systems-Theoretic Accident Model and Processes) is that system theory is a useful way to analyze accidents, particularly system accidents" (Leveson, 2004: 250). In the STAMP approach, accidents in complex systems do not simply occur due to independent component failures; rather they occur when external disturbances or dysfunctional interactions among system components are not adequately handled by the control system. Accidents therefore are not caused by a series of events but from inappropriate or inadequate control or enforcement of safety-related constraints on the development, design, and operation of the system. "Safety then can be viewed as a control problem, and safety is managed by a control structure embedded in an adaptive socio-technical system" (Leveson, 2004: 250).

STAMP is based on Rasmussen's (1997) hierarchical model of the sociotechnical system involved in risk management (see Figure 8) with control processes operating at the interfaces between vertically-adjacent levels in the hierarchy. Thus, each level in the hierarchy can be viewed as imposing constraints on the activity below it, which means that the constraints at a higher level control the behaviour at the lower level. The hierarchical model of the sociotechnical system shows a downward information flow between two hierarchical levels which imposes constraints from an upper level to a level below, and an upward feedback flow which provides the adaptive control to the complex system.

A complex system is not a static design, rather it exhibits dynamic behaviour; it is continually adapting to maintain stability and reacting to internal changes and to disturbances in its environment. The system must be designed to ensure the enforcement of constraints for safe behaviour, and its must also exhibit adaptable behaviour to cope with changes that occur during its operation. Accidents occur when the safety constraints are violated during interactions among system components. Leveson (2004) argues that, in the space shuttle *Challenger* accident the O-rings did not adequately control propellant gas release by sealing a tiny gap in the field joint, and similarly in the Mars Polar Lander loss, the software did not adequately control the descent speed of the spacecraft since it misinterpreted noise from a Hall effect sensor as an indication that the spacecraft had reached the surface of the planet. Control is also imposed by the management functions in an organisation, for example, the *Challenger* accident involved inadequate controls in the launch-decision process, and by the social and political system within which the organisation exists.

The most basic concept in STAMP is a constraint, rather than an event. Traditional accident models explain accident causation in terms of a series of events, while STAMP views accidents as the result of a lack of constraints (control laws) imposed on the system design and during operational deployment. Thus, the process that causes accidents can be understood in term of the flaws in the control loops between system

components during design, development, manufacturing, and operations. These flaws can be classified and used during accident analysis or accident prevention activities to assist in identifying all the factors involved in the accident (Leveson, 2004), and a general classification is shown in Figure 13.

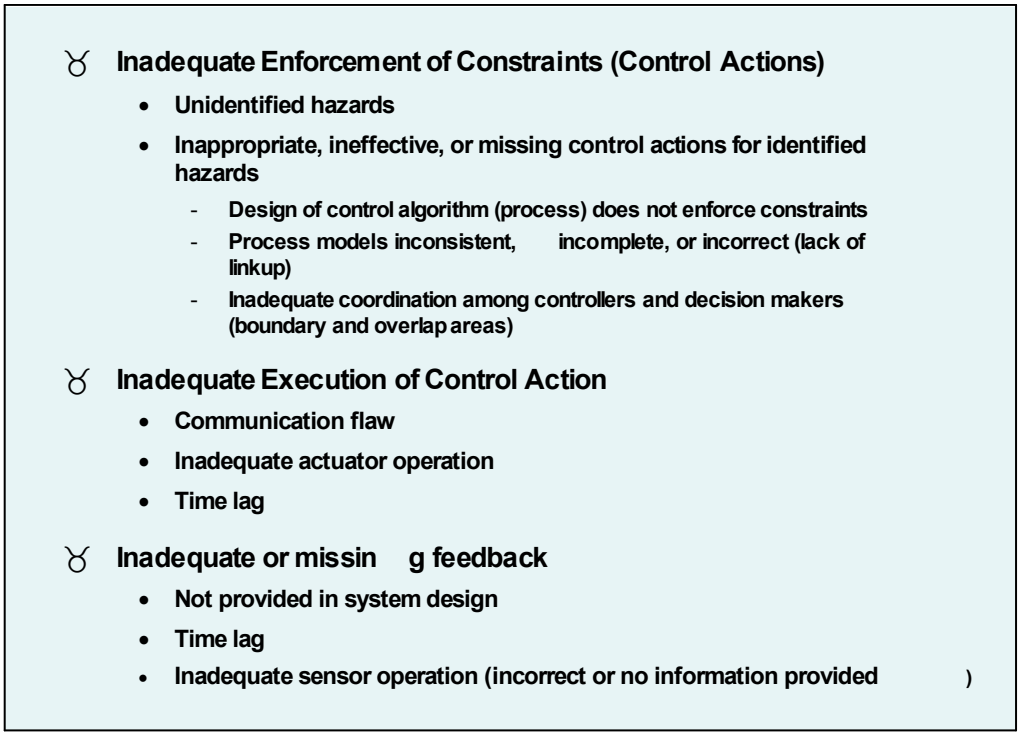

Figure 13: General Classification of Control Flaws (Leveson, 2004)

This classification investigates each control loop at every level of the sociotechnical control structure and evaluating its contribution to unsafe behaviour:

- the controller may issue inadequate or inappropriate control actions, including inadequate handling of failures or disturbances in the physical process;
- control actions may be inadequately executed, or
- there may be missing or inadequate feedback.

5.2 STAMP Analysis of the Black Hawk Fratricide

A STAMP accident analysis can be conducted in two stages:

1. Development of the Hierarchical Control Structure, which includes identification of the interactions between the system components and identification of the safety requirements and constraints;

2. Classification and Analysis of Flawed control (Constraint Failures), which includes the classification of causal factors followed by the reasons for flawed control and dysfunctional interactions.

Here we provide a summary of the STAMP analysis of the Black Hawk fratricide during the operation Provide Comfort in Iraq in 1991, which is described in detail in (Leveson et al., 2002; Leveson, 2002).

The hierarchical control structure of the Black Hawk accident is shown in Figure 14, starting from the Joint Chiefs of Staff down to the aircraft involved in the accident. At the lowest level in the control structure are the pilots who directly controlled the aircraft (operator at the sharp end).

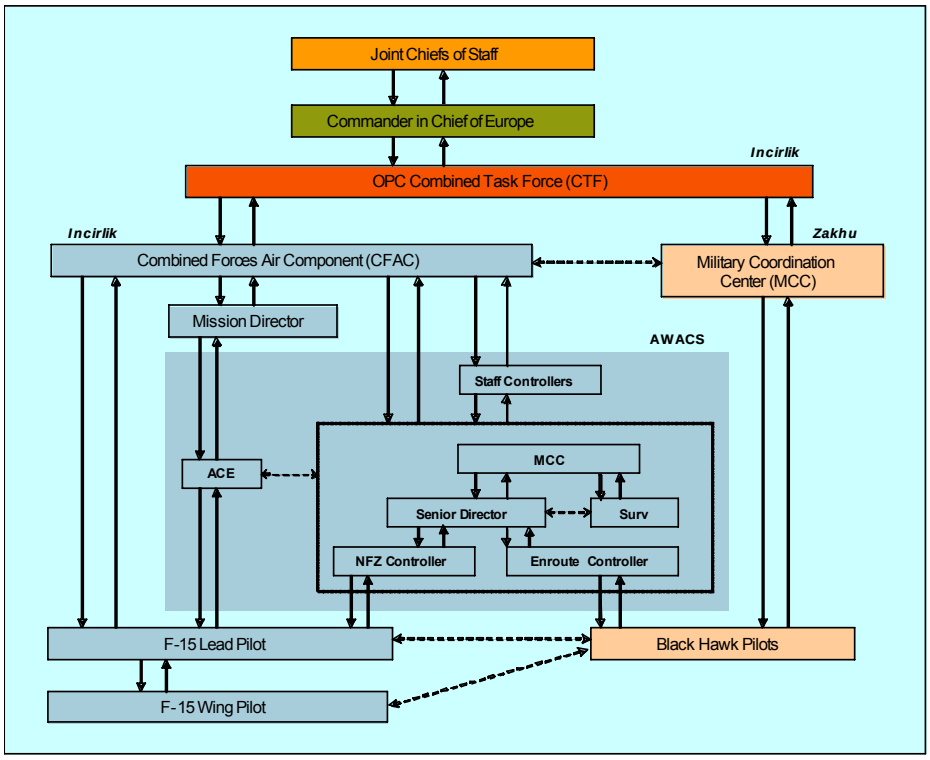

Figure 14: Hierarchical Control Structure in the Iraqi No-Fly Zone (Leveson, 2002)

The AWACS mission crew was responsible for tracking and controlling aircraft. The AWACS also carried an Airborne Command Element (ACE), who was responsible for ensuring that the larger OPC mission was completed. The ACE reported to a ground-based Mission Director. The Army headquarters (Military Coordination Center) Commander controlled the U.S. Black Hawk operations while the Combined Forces Air Component (CFAC) Commander was responsible for the conduct of OPC missions. The CFAC Commander had tactical control over all aircraft flying in the No Fly Zone (NFZ) including both Air Force fighters and Army helicopters, but had operational control only over the Air Force fixed-wing aircraft.

In addition to the formal control channels, there were also communication channels, shown in Figure 14 as dashed lines, between the process components at each level of the hierarchy.

The hierarchical control structure (Figure 14) is then analysed to identify the safety constraints at each level in the hierarchy and the reasons for the flawed control. Using the general classification in Figure 13, Leveson (2002) describes the analysis at each of the levels in the Hierarchical Control Structure. For example, at the Physical Process Level (see Figure 15), the safety constraint required that weapons must not be fired at friendly aircraft. All the physical components worked exactly as intended, except perhaps for the IFF (Identify Friend or Foe) system, which gave an intermittent response (this has never been completely explained). There were, however, several

dysfunctional interactions and communication inadequacies among the correctly operating aircraft equipment:

- The Black Hawks and F-15s were on different radio frequencies and thus could not communicate or hear the radio transmission between the two F-15 pilots and between the lead F-15 pilot and the AWACS.
- The F-15 aircraft were equipped with the latest anti-jamming HAVE-QUICK II radios while the Army helicopters were not. The F-15 pilots could have switched to non-HAVE QUICK mode to enable communication with the Black Hawk pilots; however, the procedures given to the F-15 pilots did not contain this instruction.
- The Black Hawks were not squawking the required IFF code for flying within the NFZ, and this was concluded as cause for F-15s receiving no response to their Mode IV IFF query. However, according to an Air Force analysis of the IFF system, the F-15s should have received a Mode IV response regardless of the code squawked by the targets; this contradiction has never been explained.

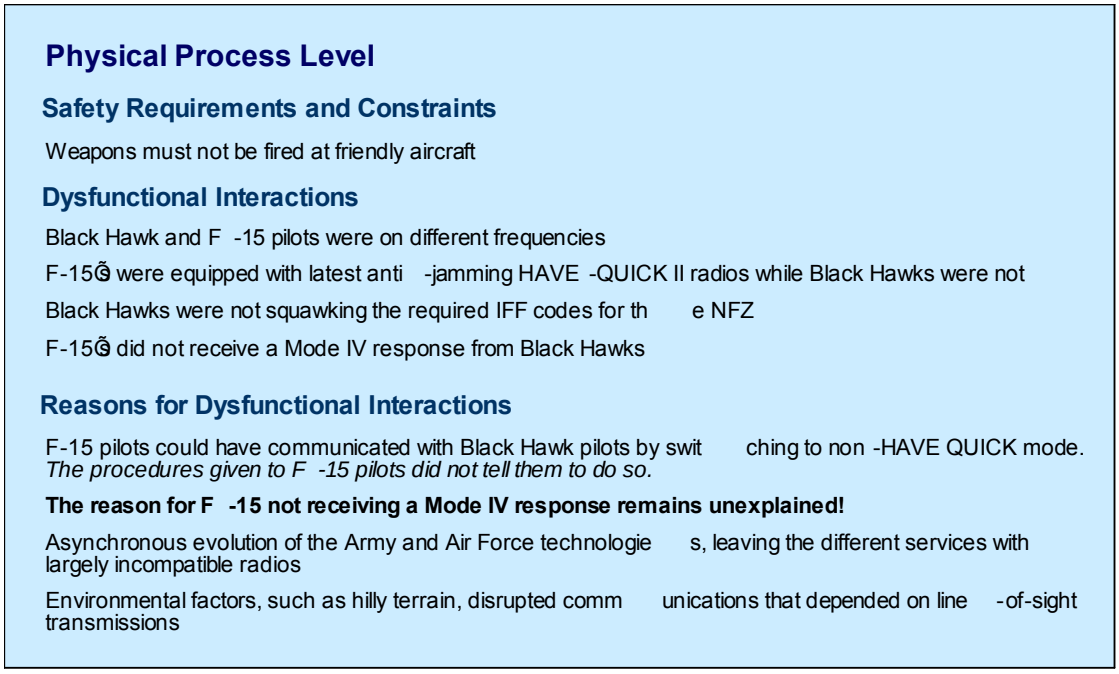

Figure 15: Physical Process Level: Classification and Analysis of Flawed Control

A major reason for these dysfunctional interactions can be attributed to the use of advanced technology by the Air Force, which was incompatible with the Army radios in the Black Hawks. The hilly terrain also contributed to the interference in the line-of-sight transmissions.

However, it is also important to analyse the safety constraints and flawed control at the higher levels in the hierarchical control structure to obtain a system-wide understanding of the contributory causal factors. Leveson (2002) conducted a detailed analysis at each of the other levels in the Hierarchical Control Structure, namely, The Pilots Level, ACE and Mission Director Level, AWACS Control Level, CFAC and MCC Level, CTF Level, and the National Command Authority and Commander-in-Chief Europe levels.

The following four causes have been generally accepted by the military community as the explanation for the shootdown (AAIB, 1994):

- The F-12 pilots misidentified the Black Hawks;
- The AWACS crew failed to intervene;
- The helicopters and their operations were not integrated into the Task Force;
- The Identification Friend or Foe (IFF) systems failed.

While there certainly were mistakes made at the pilot and AWACS levels as identified by the special Air Force Task Force and the four factors identified by the accident report were involved in the accident, the use of the STAMP analysis (Leveson, 2002) provides a much more complete explanation including:

- inconsistent, missing, or inaccurate information;
- incompatible technology;
- inadequate coordination;
- overlapping areas of control and confusion about who was responsible for what;
- a migration toward more efficient operational procedures over time without any controls and checks on the potential adaptations;
- inadequate training; and
- in general a control structure that did not enforce the safety constraints.

Leveson attributes the organisational factors at the highest levels of command for the lack of coordination and communication, as a key accident factor, which led to the failures at the lower technical and operational levels.

Using the traditional accident models based on event chains would have resulted in focusing attention on the proximate events of this accident and on the identification of the humans at the sharp end such as the pilots and the AWACS personnel. The STAMP method clearly identifies other organisational factors and actors and the role they played.

6. Formal Methods and Accident Analysis

6.1 What are Formal Methods?

A large number of approaches and methods are available for the modelling of complex and real-time computer systems, for example (Yourdan, 1989; Gomaa, 1986; Booch, 1994). Buede (2000) provides an overview of several qualitative modelling approaches used as part of the development of functional and operational architectures during the engineering of systems. These descriptions are usually given in natural languages or in diagrams, but they can be ambiguous and hard to analyse. Moreover, these techniques do not effectively capture the rich behaviour possible in a complex system, for example concurrency. Harel (1987) introduced statecharts as a generalisation of bigraphs to extend the notions of state-transition diagrams. However, statecharts still provide limited semantics and syntax for modelling complex systems.

Formal methods are mathematically-based techniques which provide a rigorous and systematic framework for the specification, design and verification of computer systems (both software and hardware). Formal methods essentially involve the use of a formal specification language composed of three main components: rules for determining the grammatical well-formedness of sentences (the syntax); rules for interpreting sentences in a precise, meaningful way within the domain considered (the semantics); and rules for inferring useful information from the specification (the proof theory) (Lamsweerde, 2000). This basis provides the means of precisely defining notations like consistency and completeness and more relevantly, specification, implementation and correctness (Wing, 1990). Formal methods are employed to model the behaviour of a system and to formally verify that the design and implementation conforms to its specification and satisfies system properties. During the early stage of system development, they can effectively identify specification and design errors. These errors might go unnoticed and discovered later either during system integration and testing or during deployment and operation, which can be very costly to fix.

Clarke & Wing (1996) provide a survey of formal methods and tools for specifying and verifying complex hardware and software systems. They assess the application of formal methods in industry and describe some successful case studies such as: the formal specification of IBM's Customer Information Control System (CICS); and, an on-line transaction processing system in the Z language; formal requirements specification for the Traffic Collision Avoidance System (TCAS) II using the Requirements State Machine Language (RSML). Model checking and theorem proving are two well-established approaches for formal verification. Clarke & Wing describe some notable examples of the successful application of these techniques and associated tools in industry.

Typically formal methods have been applied to various software development phases such as requirements analysis, specification, design and implementation (Bjørner & Druffel, 1990); they are currently mainly used for stabilising requirements and re-engineering existing systems (Gaudel, 1994; Wildman, 2002).

Wildman (2002) describes the use of formal specification techniques to reformulate the requirements of the Nulka Electronic Decoy. The Nulka Electronic Decoy is a joint Australian/US project to counter anti-ship missiles. The requirements specification contained informal natural language requirements relating both to time-related

performance requirements and to other physical characteristics. Wildman focuses on the benefits provided by the use of an appropriate formal specification notation; a dramatic improvement was achieved where 50% of the informal requirements were modified as a result of formalisation and consultation with domain experts. The formal model of requirements for the decoy Prime Item Development Specification (PIDS) aimed to:

- eliminate the difficulties associated with the *accurate* capture of these requirements in English,
- aid *understanding* of requirements, and
- provide a natural *basis* for a *clear* and *readable* English language description.

The formal analysis of the NULKA PIDS consisted of translating the original informal requirements into the Interval Calculus, and the resultant formal requirements were then manually checked for critical properties, namely, consistency, correctness, precision, and abstraction. The results of this application have demonstrated the usefulness of mathematical modelling of the English language specification and its subsequent reverse engineering back into English, which provided an *accurate*, *readable*, and *clear understanding* of the natural language specification.

The tremendous potential of formal methods has been recognised by theoreticians for a long time. There are comprehensive accounts of experience on the use of formal methods in industry and research (see for instance: Butler at al., 2002; Hinchey & Bowen, 1995). A comprehensive database of industrial and space applications is available at the Formal Methods Europe applications database (FME, 2004).

There is a variety of formal methods which support the rigorous specification, design and verification of computer systems, for example, COLD, Circal Process Algebra, Estelle, Esterel, LOTOS, Petri Nets, RAISE, SDL, VDM and Z (see for example: FMVL, 2007). Lindsay (1998) provides a tutorial example to illustrate the use of formal methods for system and software development. The example concerns part of a simplified Air Traffic Control system, using the Z notation and Cogito methodology for modelling, specification, validation and design verification.

Formal languages and methods are frequently applied to gain high confidence in the accuracy of information in the design of safety-critical systems (Hinchey & Bowen, 1995). For example, the Federal Aviation Administration's air traffic collision avoidance system (TCAS II) was specified in the formal language, RSML (Requirements State Machine Language), when it was discovered that a natural language specification could not cope with the complexity of the system (Leveson et al., 1994). Haveland & Lowry (2001) discuss an application of the finite state model checker SPIN to formally analyse a software-based multi-threaded plan execution module programmed in LISP, which is one component of NASA's Remote Agent, an artificial intelligence-based space-craft control system. A total of five previously undiscovered concurrency errors were identified in the formal model; each represented an error in the LISP code. In other words, the errors found were real and not only errors in the model. The formal verification effort had a major impact: locating errors that would probably not have been located otherwise and identifying a major design flaw. Formal approaches to development are particularly justified for systems that are complex, concurrent, quality-critical, safety and security-critical.

Formal methods presently do not scale up to the modelling and verification of large complex systems. Furthermore, there is a need for further information on the practical application of formal methods in industry to assist in the procurement, management, design, testing and certification of safety and security critical system. The formal methods group of European Workshop on Industrial Computer Systems (EWICS) have released guidelines on the use of Formal Methods in the Development and Assurance of High Integrity Systems (Anderson et al., 1998a; 1998b). These guidelines provide practical advice for those wishing to use or evaluate formal methods in an industrial environment. The employment of formal methods does not *a priori* guarantee correctness; however, they can enhance our understanding of a system by revealing inconsistencies, ambiguities, and incompletenesses that might otherwise go undetected (Clarke & Wing, 1996). Thus, the main benefits can be seen as achieving a high degree of confidence in the correctness and completeness of specifications and a high degree of assurance that the design satisfies the system specification.

6.2 The Connection between Formal Methods and Accident Analysis

The potential benefits of formal methods are well known and their use is suggested or mandated by many industry standards for the design of safety and security critical systems (Hinchey & Bowen, 1995; ATEA, 1998). However, at present, the use of formal methods is largely restricted to initial system design, and their use later in the system development lifecycle is much less common. Formal modelling of accident reports is one example of formal methods being used later in the lifecycle for *post hoc* analysis (see, for example, Ladkin & Loer, 1998; Burns, 2000).

Ensuring the quality of accident reports should be a high priority for organisations as they have a moral responsibility to prevent accident recurrence. They also have a financial responsibility to their investors; accident recurrence carries the possibility of damaging litigation and loss of customer confidence (Burns, 2000). However, the structure, content, quality, and effectiveness of accident reports have been much criticised (e.g., Burns et al., 1997; Ladkin & Loer, 1998). A large number of accident investigation reports do not accurately reflect the events, or are unable to identify critical causal factors, and sometimes conclude with incorrect causes of the accident. Omissions, ambiguities, or inaccurate information in a report can lead to unsafe system designs and misdirected legislation (Leveson, 1995). Thus, there is a critical need to improve the accuracy of the information found in conventional accident investigation reports.

Burns (2000) provides an overview of aspects of natural language accident reports that inhibit the accurate communication of the report contents:

- **Size:** The sheer size of accident reports makes it difficult for the reader to absorb all the salient points; a great deal of information can be forgotten, lost track of, or simply missed. The size also increases the chances of syntactic and semantic errors, ambiguities, and omissions in the transcription of the report.
- **Structure:** Sections of the report cannot be read in isolation, and thus the reader needs to read the full document to comprehend the information provided.
- **Validation:** Natural language has no accepted formal syntax or semantics and so it is not currently possible for any mathematical analysis to be conducted over a natural language report. The informal process of peer review and verbal

descriptions does not ensure high quality in the argumentation; consistency and coherence errors are still found in the validity of the reasoning in accident reports. For example, managerial wrongdoing is highlighted as a causal factor in the Challenger Shuttle crash, yet the report fails to address the question of why the management ignored the advice of their engineers and made the decision to launch (Vaughn, 1996).

- **Differing Viewpoints:** Every individual involved in the accident and investigation has a view on what happened in the accident and why. Reports are written by a number of authors, and the information presented in particular sections will be affected by the author's mental representation of the accident. Usually the scientific evidence presented demonstrates a considerable divergence of views.

- **Redundancy:** Repetition is common in reports: events may be summarised, and then described in more detail; the same events may be described from a different viewpoint in different chapters. Coherence and consistency therefore become an issue. Redundant detail increases the overall size of the report and there is no reliable means to check consistency between similar information. As with some of the other weaknesses we highlight, redundancy is not always bad. In a large document, some repetition does help the reader.

- **Imprecision of Natural Language:** Sentences and terms in natural language may have more than one interpretation. The context of the sentence or common sense of the reader is often relied upon to extract the intended interpretation, but these are not always sufficient. If the semantics of a statement cannot be uniquely determined, the accuracy of the information communicated to the reader by the report cannot be assured. It is important that any two people (e.g., the report writer and a system designer) do not interpret an accident report differently.

- **Representing Concurrency:** In a complex system, many different events can occur simultaneously. It can be important, during the analysis of the accident, to be aware of what is happening in a particular time interval. Natural language is poorly suited to represent such concurrency.

- **Distinguishing Prescriptive and Descriptive Behaviour:** The behaviour of the agents involved in an accident is unlikely to have been ideal or even expected. However, accident reports are generally not written to apportion blame, and this can make it difficult to ascertain from the text when certain behaviour is acceptable and when it is not.

- **Incompleteness:** The scope of the accident can be defined as the chronological range between the first and last relevant event in the accident scenario. The separate viewpoints in each chapter can make the scope of the accident unclear. If the reader misinterprets the scope, their visualisation of the accident will be incomplete and inaccurate.

- **Politics of Inquiries:** Depending on the nature of the accident, a public investigation will be conducted under the provision of different legislative acts; for example, in Australia an accident involving a train would be investigated under the Regulations of Railways Act. An accident involving a commercial aeroplane would be investigated under the Civil Aviation (Investigation of

Accidents) Regulations. Where there is no specific legislation, the Health and Safety at Work Act can initiate formal investigations. Inconsistency caused by the numerous separate mechanisms for instigating investigations can make the quality and content of reports extremely variable.

Formal methods can improve accident analysis by emphasising the importance of precision in definitions and descriptions, and providing notations for describing and reasoning about certain aspects of accidents. Ladkin & Loer (1998) describe a formal method, called *Why-Because Analysis*, for accident modelling and rigorous reasoning; and have demonstrated benefits in the application of this method to a number of case studies in aviation and train accidents (for example: Höhl & Ladkin, 1997; Ladkin, 2005). The development of deontic action logic as a language for constructing formal models (Burns, 2000) has demonstrated that the methodical construction of a formal model of the accident report prevents, improves, or makes explicit weaknesses in natural language reports, in particular the accuracy and presentation of the information with accident reports.

6.3 Logic Formalisms to Support Accident Analysis

During the last decade many attempts have been made on the use of formal methods for building mathematically-based models to conduct accident analysis. During the last decade in particular, many attempts have been made on the use of formal methods for building mathematically-based models to conduct accident analysis. A comprehensive survey on the application of various formal logics and techniques to model and reason about accident causation is given by Johnson & Holloway (2003a) and Johnson (2003). They discuss the weakness of classical (propositional) logic in capturing the different forms of causal reasoning that are used in accident analysis. In addition, the social and political aspects in accident analysis cannot easily be reconciled with the classical logic-based approach. Johnson & Holloway argue that the traditional theorem proving mechanisms cannot accurately capture the wealth of inductive, deductive and statistical forms of inference that investigators routinely use in their analysis of adverse events.

Thomas (1994) used a first order logic to formalise the software code known to be a source of error in the Therac-25 radiation machine (Leveson, 1993). The automated theorem prover LP (Larch Prover) was employed to reason about the behaviour of the code, which helped identify the underlying cause of the unexpected behaviour of the code that contributed to the accident. This approach assisted in correcting the software error and in providing rigorous evidence (via formal proofs) that the modified software executed according to the expected/specified behaviour.

Fields et al. (1995) employed CSP Process Algebra (Hoare, 1985) to formally specify both the tasks of the human operators and the behaviour of the system. The performance model contributed to the analysis of human error in system failures by identifying the sequence of actions (erroneous traces) related to the failure modes of the operator.

Petri nets have been successfully used for dynamic modelling of parallel and concurrent systems with time constraints in a wide range of applications including safety-critical systems. Vernez et al. (2003) provide a review of the current uses of Petri nets in the fields of risk analysis and accident modelling. They demonstrate that Petri nets can explicitly model the complex cause to consequence relationships between

events. Vernez et al. provide a translation of key safety concepts onto the Petri nets formalism and suggest that this can facilitate the development of accident models. They investigate the modelling capability of Coloured Petri Nets (CPN) to predict possible accident scenarios in the Swiss metro, a high-speed underground train planned for interurban linking in Switzerland. Relevant actors, events and causal relationships were translated into the CPN formalism, and the Design/CPN tool and the state space method were employed to analyse the states or accident scenarios (a succession of possible system states) generated in the occurrence graph. Vernez et al. argue that the results obtained in the CPN modelling and analyses of accident processes are realistic as compared to both previous tunnel accidents and tunnel safety principles.

Burns et al. (1997) applied a Sorted First Order Logic (SOFAL) to specify and reason about the human contribution to major accidents. SOFAL has the advantage over other formalisms such as first order logic and Petri nets, that it explicitly specifies agents (people, operators) as distinguished from other system objects (such as inanimate objects). This feature supports the analysis of accidents in focusing more on those objects which directly affect the system behaviour. SOFAL can also support reasoning over temporal aspects of the system behaviour (Burns et al., 1997), for example it can demonstrate that there exists a sequence of actions which, when performed, will lead to a scenario where an accident occurs.

Deontic logics were developed for reasoning about norms in complex organisational and procedural structures within a system, in particular, to reason about notions in ethics and philosophy of laws (Wieringa & Meyer, 1994). Deontic logic can also express normative (e.g. legal) and non-normative (e.g. illegal, non-permitted) behaviour, which can be used for modelling and reasoning about ideal (normative) and non-ideal (non-normative) or actual system behaviour which is commonly found in accidents (Burns, 2000). The concept of non-normative scenarios is important in accidents as it can be used to model some non-ideal, non-permitted or illegal behaviour (such as smoking in a “non-smoking” zone), which if occurs may lead to an accident. Burns (2000) describes an Extended Deontic Action Logic (EDAL) language for formally modelling accident reports, which is an extension of Deontic Action Logic (Khosla, 1988). EDAL models both the prescribed (expected) and the actual behaviour of the system and the relationship between the two; this facilitates an analysis of the conflict between the two behaviours and thus can greatly assist in understanding the causal factors in an accident. Using the Channel Tunnel fire accident report as a case study, Burns (2000) developed a formal model in EDAL and demonstrated that this approach can be used to reason about qualitative failure, errors of omission and commission, and prescriptive failures. For example, EDAL enabled the specification and analysis of where, how, and by whom, norms were broken within the system. Burns has demonstrated that constructing and reasoning about formal accident report models highlights problems in the accident report.

The focus of EDAL has been on the deontic modalities in an accident, other formal modelling techniques, such as Why-Because Analysis (Ladin & Loer, 1998), have considered further aspects of accidents such as epistemic and real-time behaviour.

6.4 Probabilistic Models of Causality

The accident modelling approaches discussed so far are based on deterministic models of causality. These models focus on the identification of deterministic sequence of cause and effect relationships, which are difficult to validate (Johnson, 2000). For example, it cannot be guaranteed that a set of effects will be produced even if necessary and sufficient conditions can be demonstrated to hold at a particular moment. Johnson argues that the focus should be on those conditions that make effects more likely within a given context, and examines the application of probabilistic models of causality to support accident analysis. Probabilistic causation designates a group of philosophical theories that aim to characterise the relationship between cause and effect using the tools of probability theory (Hitchcock, 2002). The central idea underlying these theories is that causes raise the probabilities of their effects. Johnson proposes an approach for the causal analysis of adverse accidents that is based on the integration of deterministic and probabilistic models of causality.

The use of conditional probabilities has some significant benefits for accident analysis (Johnson, 2000); for example, in the Nantichoke fire we need to know the probability of ignition from each source (indicator taps, exposed manifolds) given the fuel leak characteristics. Johnson & Holloway (2003a) discuss the use of Bayesian Logic (which exploits conditional probabilities) for accident analysis, as an example to reason about the manner in which the observation of evidence affects our belief in causal hypothesis.

The probabilistic theory of causality has been developed in slightly different ways by many authors. Hitchcock (2002) conducts a review of these developments and discusses the issues and criticism to the probabilistic theories of causation. Here, we discuss the mathematical theory of causality developed by Pearl (2000), which is a structural model approach evolved from the area of Bayesian networks. The main idea behind the structure-based causal models is that the world is modelled by random variables, which may have causal influence on each other (Eiter & Lukasiewicz, 2001). The variables are divided into exogenous variables, which are influenced by factors outside the model, and endogenous variables, which are influenced by exogenous and endogenous variables. This latter influence is expressed through functional relationships (described by structural equations) between them.

Formally, Pearl (2000) defines a **causal model** as a triple $M = (U, V, F)$ where:

(i) U is a set of *background variables*, (also called *exogenous*), that are determined by factors outside the model;

(ii) V is a set $\{V_1, \ldots, V_n\}$ of variables, called *endogenous*, that are determined by variables in the model - that is, variables in $U \cup V$ and

(iii) F is a set of functions $\{f_1, f_2, \ldots, f_n\}$ such that each f_i is a mapping from (the respective domains of) $U \cup (V \setminus V_i)$ to V_i and such that the entire set F forms a mapping from U to V. In other words, each f_i tells us the value of V_i given the values of all other variables in $U \cup V$, and the entire set F has a unique solution $V(u)$. Symbolically, the set of functions F can be represented by writing: $v_i = f_i(pa_i; u_i)$, $i = 1, \ldots, n$,
where, pa_i is any realization of the unique minimal set of variables PA_i in $V \setminus V_i$ (connoting parents) sufficient for representing f_i. Likewise, $U_i \subseteq U$ stands for the unique minimal set of variables in U sufficient for representing f_i.

The relationship between the variables of a causal model $M = (U, V, F)$ can be associated with the *causal graph* for M, which is the directed graph that has $U \cup V$ as the set of nodes and the directed edges point from members of PA_i and U_i towards V_i (Pearl, 2000). This graph merely identifies the endogenous and background variables that have direct influence on each V_i; it does not specify the functional form of f_i.

Pearl (2000) uses the structural causal model semantics and defines a **probabilistic causal model** as a pair $(M, P(u))$ where M is a causal model and $P(u)$ is a probability function defined over the domain of the background variables U.

Pearl (2000) has also demonstrated how counterfactual queries, both deterministic and probabilistic, can be answered formally using structural model semantics. He also compares the structural models with other models of causality and counterfactuals, most notably those based on Lewis's closest-world semantics.

A number of research groups are investigating the use, extension and development of formal languages and methods for accident modelling and analysis, such as the Glasgow Accident Analysis Group (GAAG, 2006) and the NASA Langley formal methods research program on accident analysis (LaRC, 2004). The research program at NASA Langley is investigating the suitability of using one or more existing mathematical representations of causality as the basis for developing tools for:

- explaining causes and contributing factors to accidents;
- analysing causal explanations for consistency, completeness, and other desired characteristics;
- storing causal explanations for retrieval; and
- using previously stored causal explanations in the design of new systems.

Formal methods have been applied successfully to the design and verification of safety-critical systems; however, they need to be extended to capture the many factors and aspects that are found in accidents and accident reports. A single modelling language is unlikely to model all the factors and aspects in an accident (Burns, 2000). Also scaling up, formal methods have limitations to model complete sociotechnical systems, they need specialists in mathematics, and not everything can be formalised.

6.5 Why-Because Analysis (WBA)

6.5.1 WBA Method

Why-Because Analysis (WBA) is a method for the failure analysis of complex, open, heterogeneous systems (Ladkin, 1999). The adjective "open" means that the behaviour of the system is highly affected by its environment, and "heterogeneous" refers to a system comprised of a group of closely connected components that are not alike, such as digital, physical, human and procedural components, which are all supposed to work together. For example, modern aviation operations have all of these components and thus form a complex, open, heterogeneous system.

The investigation of failures of complex systems is a wide field of practical interest that has traditionally not been carried out with any significant use of formal methods. Ladkin & Loer (1998) developed the formal Why-Because Analysis (WBA) method, which enables one to develop, then formally to prove, the correctness and relatively

sufficiency of causal explanations. This formal technique is based on formal semantics and logic, and separates the various explanatory domains: time, causation, and deontics (regulations, obligations and operating procedures). The primary application domain to date has been transportation accidents, concentrating mainly on aircraft accidents. WBA stems from an initiative to increase the objectivity of accident investigations by encouraging "rigorous causal analysis". WBA is primarily concerned with analysing causality, and allows objective evaluation of events and states as causal factors.

In general, the term "cause" is not well defined and there is little consensus on what constitutes a cause. One philosophical approach to causation views counterfactual dependence as the key to the explanation of causal facts: for example, events *c* (the cause) and *e* (the effect) both occur, but had *c* not occurred, *e* would not have occurred either (Collins et al., 2004). The term ''counterfactual'' or ''contrary-to-fact'' conditional carries the suggestion that the antecedent of such a conditional is false.

David Lewis (1973) developed a number of logics to capture counter-factual arguments that provide a formal semantics for causation. Lewis's semantics can be used to state that *A* is a causal factor of *B* (where *A* and *B* are two events or states), if and only if *A* and *B* both occurred and in the nearest possible worlds in which *A* did not happen neither did *B*. This implies that *A* is a cause of *B* or *A* is a necessary causal factor of *B*. Lewis's theory employs a possible world semantics for counterfactuals, where the central notion is a relation of *comparative similarity* between worlds. One world *W'* is considered to be *closer to the actual world W* than another world *W"* if *W'* resembles the actual world more than *W"*. Lewis's semantics require one to consider the *nearest possible worlds to the actual world* in which *A* did not occur, and to consider whether *B* did not occur in these situations either. Ladkin (1999) explains the notion of *nearest* possible worlds intuitively by using an example where his office door is open. "The nearest possible worlds in which my door is shut is one in which my door is shut, air currents around it behave appropriately, sound through it is muffled as it should be, but broadly speaking everything else remains the same. A further-away world would be one in which someone else who is not me is sitting here typing, and even further-away world is one in which this whole environment is situated in Ghana rather than Germany." It precludes the observation that other causal factors may have led to the shut door in any of the possible worlds. However, this does not rule out the existence of alternative causes: it implies that those causes may only arise in worlds that are remote from the present one that is under consideration (Ladkin & Loer, 1998**).**

Ladkin & Loer (1998) introduce notations and inference rules which allows them to reduce the Lewis criterion for counterfactuals in the form (Figure 16). This logic provides semantics for informal concepts such as "cause" that is used to explain the causal-factor relation between facts *A* and *B*.

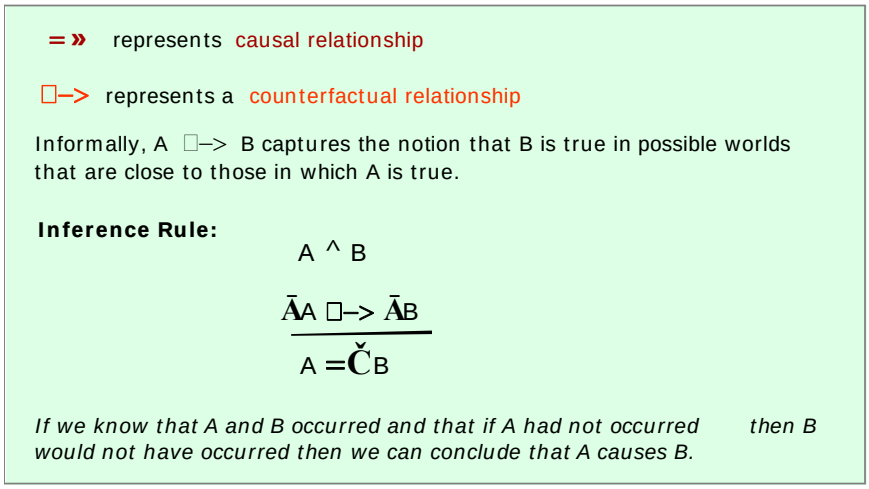

Figure 16: WBA formal notations and rules for causal relation

Lewis's semantics for causation in terms of counterfactuals, and the combination of Lamport's Temporal Logic (Lamport, 1994) and other logics into a formal logic called Explanatory Logic, form the basis of the formal method WBA. WBA is based around two complementary stages:

1) Construction of the WB-Graph; and
2) Formal Proof of Correctness of the WB-Graph

WBA begins with a reconstruction phase, where a semi-formal graphical notation models the sequences of events leading to an accident. The significant events and states are derived from the accident investigation report in their proper time order. These sequences can be represented in a form of temporal logics and then each pair is iteratively analysed to move towards a causal explanation using Lewis's counterfactual test. The graph of this relation is called a WB-Graph (see Figure 19 as an example).

The WB-Graph is subjected to a rigorous proof to verify that: the causal relations in the graph are correct, that is they satisfy the semantics of causation defined by Lewis; and there is a sufficient causal explanation for each identified fact that is not itself a root cause. The formal logics employed in the WBA formal proof method are shown in Figure 17. A detailed development of the formal proof of correctness and the EL logic is described in Ladkin & Loer (1998).

Method	Used for:
modal logic/Tense logic	- temporal reasoning
Lewis counterfactuals	- causal explanation
Deontic reasoning	- incorporation of Standard Operating Procedures (SOP) - incorporation of regulatory environment - incorporation of significant non -events - reasoning about latent errors

Figure 17: WBA Logics

The WBA method has been used for analysing a fairly large number of accident reports, mainly for aircraft accidents. In the Lufthansa A320 accident in Warsaw, the logic of the braking system was considered the main cause of the accident. The accident report contained facts that were significantly causally related to the accident. However, these facts were not identified in the list of "probable cause/contributing factors" of the accident report.

WHY	BECAUSE	DESCRIPTION
[0]		accident
	/\ [1]	death of 1st person
	/\ [2]	death of 2nd person
	/\ [3]	damage to AC
[1]	[3.1]	(AC hits earth bank)
[2]	[-.1]	asphyxiation
	[2.1] /\ <-.1>	smoke in cabin
	/\ [-.2]	remained in cabin
	<2.1.1> [3.2]	(AC burns)
	[2.1.2] /\ <-.1>	unconsciousness
	/\ [<-.2>]	unnoticed during evacuation
	<2.1.2.1> [3.1]	(AC hits earth bank)
	[<2.1.2.2>] [<-.1>]	motionless, noiseless, position, smoke in cabin, time pressure, etc.
[3]	/\ [-.1]	AC hits earth bank
	/\ <-.2>	AC burns
	[3.1] /\ [-.1]	AC overruns RWY
	/\ <-.2>	earth bank in overrun path
	[3.1.1] /\ [<-.1>]	certain cause: excessive speed on landing
	/\ <-.2>	certain cause: unstabilised approach
	/\ [<-.3>]	certain cause: braking delayed

Figure 18: Extract of Textual Form of WB-Graph from the Warsaw Accident (Höhl & Ladkin, 1997)

Höhl & Ladkin (1997) analysed the text of the accident report and identified the relevant states and events concerning the accident. The events and states were used to prepare a textual version of the WB-Graph with path numbering (Figure 18). The WB-Graph commences from the accident event (node 0), and proceeds via a backwards-chronological search investigating which events and states were causal factors. This search continues with reasoning about why each subsequent event occurred until a source events or state is reached. A source node has no incoming links i.e. they have no significant causal factors and are considered as the original source to a sequence of events, such as nodes *3.1.2*, *3.1.1.3.2.2* and *3.1.1.3.1.1.1.3* in Figure 19. The information and path numbering in the textual version is then used to draw the WB-Graph (Figure 19). The WB-Graph can be used to answer questions such as, Why did an event X happen? The event X happened because of the events A, B and C. The "Because part " shows the conjunction of explanations (events A, B, C) why the event X happened. The causal graph grows by investigating why the next event, such as X.A happen, and

explaining that X.A occurred because of events D and F. Therefore, the event 3.1, Aircraft hits earth bank, occurred because of event 3.1.1, Aircraft overruns runway, and state 3.1.2, earth bank in overrun path. The state 3.1.2 occurred because of the source node 3.1.2.1, earth bank was built by airport authority for radio equipment.

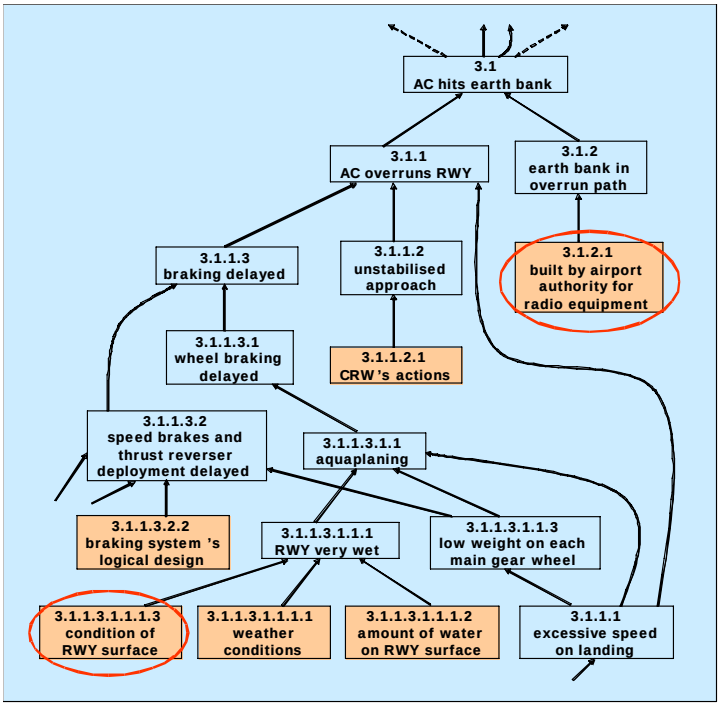

Figure 19: WB-Graph Extract of the Warsaw Accident (Höhl & Ladkin, 1997)

The rigorous reasoning employed in the WBA-Method enabled Höhl & Ladkin (1997) to identify two fundamental causes (source nodes in the WB-graph) that occurred in the accident report but were omitted as "probable cause" or "contributing factors": the position of the earth bank (node 3.1.2.1), and the runway surfacing (node 3.1.1.3.1.1.1.3). Once the position of the earth bank was identified as an original causal factor, it can be concluded that had the bank not been where it is, the accident would not have happened. Also, if the condition of the runway surfacing had been otherwise, the wheel braking systems could have functioned earlier and perhaps the collision with the bank avoided. The rigorous reasoning in the WB-Graph enabled the recommendation of appropriate preventative strategies, e.g., removal of earth bank to provide a free overrun area, and to mitigate the occurrence of future similar accidents.

Thus the WB-Graph helped to identify logical mistakes in the accident report. This example has illustrated how the WB-method renders reasoning rigorous, and enables the true original causal factors to be identified from amongst all the causally-relevant states and events.

6.5.2 WBA of the Black Hawk Fratricide

Twenty-six people died by friendly fire during peace-keeping operations after the Gulf War on April 14, 1994, when two U.S. Air Force F-15 fighters shot down two U.S. Army Black Hawk helicopters in the no-fly zone over northern Iraq (AAIB, 1994; GAO, 1997). The major reasons for this fratricide are attributed to multiple coordination failures at the individual, group and organisational levels in a complex command and control

structure (see Figure 14). It is interesting to note that there were no notable technical failures; in fact the failure of the IFF system in the F-15s to receive the Black Hawks' identification code remains unexplained.

Snook (2000) employed social and organisational theories to explain the accidental shootdown of the two Black Hawk helicopters. He developed a timeline of significant events and a complex Causal Map of the incident. Ladkin & Stuphorn (2003) conducted a Why-Because Analysis of facts as presented in the Executive Summary of the U.S.A.F. Aircraft Accident Investigation Board report (AAIB, 1994), and compared their analysis with Snook's Causal Map.

Event	*Time (Zulu)*	*F15s (Tiger)*	*Black Hawks (Eagle)*	*AWACS (Cougar)*
Departure Incirlik	0436			X
Departure Diyarbakir	0522		X	
On Station	0545			X
.H. displayed on SD radar scope	0612			X
Radio transmission at .Gate.	0621		X	R
Eagle Track annotated .EE01.				X
Eagle land at Zakhu	~0624		X	
Eagle IFF and Radar fade			X	X
Departure Incirlik AB	0635	X		
Tiger IFF Mode IV interogated	0636	X		X
Onroute from Zakhu to Irbil	0654			X
Radio call received			X	R
.EE01. reinitiated				X
.H. regularly displayed	0655		X	X
Check In	0705	X		R
.H. ceases to be displayed	0711		X	X
Eagle enter mountainous terrain	0712			X
Eagle radar and IFF fade			X	X
Symbology continues at last known speed and direction				X
ASO places SD scope in vincinity of *Eagle* last known position	0713			X

Figure 20: Partial Time Line of Significant Events (Ladkin & Stuphorn, 2003)

A timeline is generally considered useful, in which actors are represented along with the times of events in which they participated. Ladkin & Stuphorn identified a number of ambiguities in the method used by Snook to develop the timeline of significant events. They constructed a single vertical timeline of all events, and annotated the events with the actors participating in this event, as shown in Figure 20. Thin columns lying to the right of the time line represent the actors, and a mark (a cross) in a column by an event indicates that the corresponding actor participated in that event. Use of a vertical timeline with columns for actor participation allows easily for a greater number of actors than appears visually feasible using Snook's representation.

0. Loss of 2 Black Hawk Helicopters & 26 people
1. Operation Provide Comfort
   a) directed in April 1991 by US National Command Authority
   b) Tactical Area of Responsibility north of 36 degrees lat, Iraq
2. OPORD 004 (14. Sept. 1991)
   a) Withdrawal of OPC Battalion Task Force
   b) Increase size of CTF air forces
   c) retention of the JSOTF at Incirlik AB
3. OPLAN 91-7 provided comprehensive guidance for OPC in July 91
4. Redeployment of Battalion TaskForce in Sept. 1991 (2a)
5. OPLAN 91-7 not updated

●●●

73. Shootdown of Eagle Lead (AAIB Timeline)
74. Shootdown of Eagle Trail (AAIB Timeline)
75. F15 lead fires AIM120 at Eagle Trail (AAIB Timeline)
76. F15 wing fires AIM9 at Eagle Lead (AAIB Timeline)
77. Eagle Flight used code for outside TAOR (AAIB ExSum Vol1)

Figure 21: A Partial List of Facts (Adapted from: Ladkin & Stuphorn, 2003)

Ladkin & Stuphorn (2003) derived the List of Facts (a partial list is shown in Figure 21) directly from the AAIB report (AAIB, 1994). This list of Facts differs considerably from that of Snook, and Ladkin & Stuphorn argue that the nodes in Snook's Causal Map do not appear to correspond to the facts in the AAIB report. They conducted a methodological check of Snook's Causal Map by checking the relations of the nodes to each other using the Counterfactual Test. Ladkin & Stuphorn concluded that one-quarter of the causal connections proposed by Snook were not correct since they did not pass the Counterfactual Test.

A WB-Graph was constructed from the List of Facts, and the Counterfactual Test was applied to determine the *necessary causal factor* relation amongst them (Ladkin & Stuphorn, 2003). The complete WB-Graph that was produced is quite hard to read, and has been split into three parts, the top part, the middle section, and the lower part; the top part of the WB-graphs is reproduced in Figure 22 showing the links to the two lower parts. The WB-Graphs illustrate the accuracy in the causal explanation and the advantages of applying a methodological approach such a WBA to the task of determining causality.

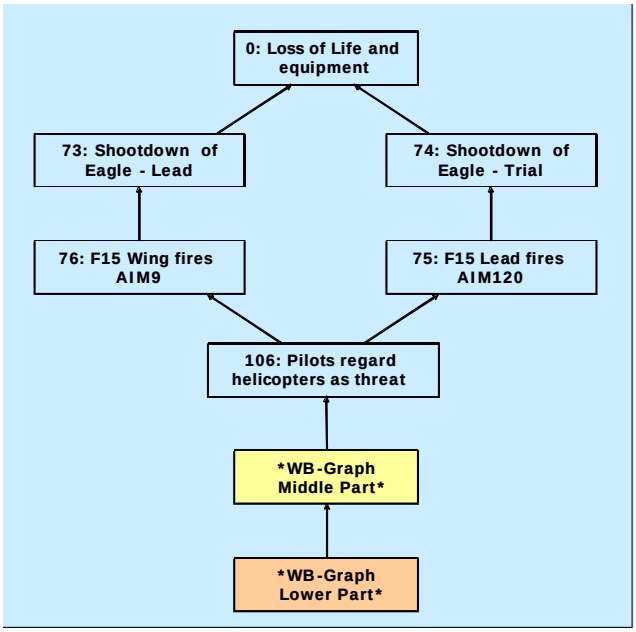

Figure 22: The AAIB WB-Graph, Top Part (Ladkin & Stuphorn, 2003)

7. Sociological and Organisational Analysis of Accident Causation

7.1 Sociological and Organisational Perspective

Major accidents such as Bhopal and Challenger have highlighted the fact that in seeking the causes of complex system accidents we must now consider the interaction and interdependence between technological and organisational systems. Shrivastava (1992) argues that industrial accidents have identifiable causes, namely, human, organisational, and technological, and their consequences demand new policies designed to prevent such crises in the future. Bhopal is only one dramatic example of how the rapid and haphazard infusion of new, sophisticated technologies put stress on the economic and social infrastructure of a community. Severe stress turns industrial accidents, such as Bhopal, into crises that lead to a pervasive disintegration in the social organisation.

A number of studies on aviation and maritime accidents have shown the human and organisational factors as major contributors to accidents and incidents. Johnson & Holloway (2007) analysed major aviation and maritime accidents in North America during 1996-2006, and concluded that the proportion of causal and contributory factors related to organisational issues exceed those due to human error. For example, the combined causal and contributory factors of aviation accidents in the USA showed: 48% related to organisational factors, 37% to human factors, 12% to equipment and 3% to other causes; and the analysis of maritime accidents classified the causal and contributory factors as: 53% due to organisational factors, 24-29% as human error, 10-19% to equipment failures, and 2-4% as other causes.

Hopkins (2000) examined the findings of the Royal Commission, from a cultural and organisational perspective, into the Esso gas plant explosion at Longford, Victoria in September 1998. This accident resulted in the death of two workers, injured eight others and cut Melbourne's gas supply for two weeks. Hopkins argues that the accident's major contributory factors were related to a series of organisational failures: the failure to respond to clear warning signs, communication problems, lack of attention to major hazards, superficial auditing and, a failure to learn from previous experience. The Royal Commission in Australia invited Hopkins as an expert witness to the Longford inquiry, he looked at this with astonishment and remarked that, "It is most unusual in this country for a sociologist to be called as an expert witness in a disaster or coronial inquiry, but in accepting my evidence the Commission was acknowledging the value of the sociological approach to its inquiry" (Hopkins, 2000: Preface).

Hopkins was a member of the Board of Inquiry into the F-111 chemical exposure of RAAF maintenance workers (Clarkson et al., 2001). He identified many cultural and organisational causes of this incident, and employed the AcciMap technique to produce a diagram identifying the network of causes that contributed to the damage done to the health of the Air Force workers (see Figure 12). Hopkins (2005) discusses various aspects of the Air Force culture and identified several fundamental values which contributed to the incapacity of the Air Force to recognise and respond to what was happening to its fuel tank workers. This emphasises the significance of

organisational factors and their influence to safety in the workplace (see also, Blackman et al., 2000).

NASA's Space Shuttle Challenger disintegrated in a ball of fire 73 seconds after launch on 28 January 1986. The Rogers Commission Report (1986) on the Space Shuttle Challenger Accident identified the cause of the disaster: the O-rings that seal the Solid Rocket Booster joints failed to seal, allowing hot gases at ignition to erode the O-rings, penetrate the wall of the booster, which finally destroyed *Challenger* and its crew. The Commission also discovered an organisational failure in NASA. In a midnight hour teleconference on the eve of the *Challenger* launch, NASA managers had proceeded with launch despite the objections of contractor engineers who were concerned about the effect of predicted cold temperatures on the rubber-like O-rings. Further, the investigation discovered that NASA managers had suppressed information about the teleconference controversy, violating rules about passing information to their superiors; NASA had been incurring O-ring damage on shuttle missions for years. The Rogers Commission Report also identified "flawed decision making" as a contributing cause of the accident, in addition to other causal factors such as production and schedule pressures, and violation of internal rules and procedures in order to launch on time.

Vaughn (1996) rejects the prevalent explanations (provided by traditional safety engineering techniques) of the cause of the *Challenger* shuttle accident and presents an alternative sociological explanation that explores much deeper cause of the failure. Vaughan discusses how common errors and violation of procedures can be seen as a normal occurrence, a concept known as normalisation of deviance. She reveals how and why NASA decision makers, when repeatedly faced with evidence that something was wrong, normalised the deviance so that it became acceptable to them. She identified three major elements behind the Challenger accident:

- An enacted work group culture, that is how culture is created as people interact in work groups;
- A culture of production built from occupational, organisational, and institutional influences; and
- A structure induced dispersion of data that made information more like a body of secrets than a body of knowledge, which silenced people.

These elements had shaped shuttle decision making for over a decade. What was unique in this particular situation was that this was the first time all three influences came together simultaneously across multiple levels of authority and were focused on a single decision to meet the Challenger launch deadline.

The physical cause of the loss of *Columbia* and its crew was a breach in the Thermal Protection System on the leading edge of the left wing, caused by a piece of insulating foam which struck the wing (CAIB, 2003). The foam impact caused a crack in the wing that allowed superheated gas to penetrate through the leading edge insulation and progressively melt the aluminium structure of the left wing, resulting in the disintegration of the Orbiter during re-entry on 1st February 2003.

The Columbia Accident Investigation Board reviewed the contemporary social science literature on accidents and invited experts in sociology and organisational theory. These experts examined NASA's organisational, historical and cultural factors and provided insights into how these factors contributed to the accident (CAIB, 2003). In

the Board's view, NASA's organisational structure and culture was equally a causal factor of the accident as the physical cause (the foam debris strike). In particular, Vaughan recognised similarities between Columbia and Challenger accidents in that both accidents resulted due to organisational system failures, and presented a causal explanation that links the culture of production, the normalisation of deviance, and structural secrecy in NASA. (CAIB 2003: Chap. 8).

Complex technological systems have many interrelated parts, and component failures in one or more parts of the system interact in unanticipated ways that lead to catastrophic accidents. Organisations managing and operating high-risk technologies can be considered as complex sociotechnical systems with systemic dependencies and tight coupling in the organisation structure and management policies, which can lead to organisational failures as contributory causal factors in system accidents. It is important to consider the organisational context in which such technological systems operate as it adds to their complexity and susceptibility to the occurrence of system accidents. Sagan (1993) examines two important schools of thought in organisation theory, namely, Normal Accident Theory and High Reliability Organisation theory, concerning the issue of safety and reliability of organisations involved in the development, management and operation of complex technological systems such as nuclear power plants, petrochemical plants, and nuclear weapons. Sagan argues that organisation theories on accidents and risk are necessary to understand and address the social causes of an accident and in enhancing performance in technologically complex organisations to safely operate and manage high-risk technological systems.

7.2 Safety Culture

The *Columbia* investigation Report identifies a "broken safety culture" as a focal point of the accident's organisational causes (CAIB, 2003). The report examines how NASA's organisational culture and structure weakened the safety structure that created structural secrecy, causing decision makers to miss the threat posed by successive events of foam debris strikes. Organisational culture refers to the values, norms, beliefs, and practices that govern how an institution functions. Schein (1992) refers to shared basic assumptions and provides a more formal definition as follows:

> *The culture of a group can now be defined as a pattern of shared basic assumptions that the group learned as it solved its problems of external adaptation and internal integration, that has worked well enough to be considered valid, and therefore, to be taught to new members as the correct way to perceive, think, and feel in relation to those problems (Schein, 1992: 12).*

A safety culture is the set of assumptions, and their associated practices, through which a group understands and conceives the dangers and hazards of the world that should be minimised for exposure to people and society (Pidgeon, 1991). Pidgeon discusses that a safety culture is created and recreated as members of a group repeatedly behave and communicate in ways which seem to them to be "natural", obvious and unquestionable, and as such will serve to construct a particular version of risk, danger and safety. Organisational culture has an influence on the overall safety, reliability and effectiveness of the operations in an organisation. Safety culture is a part of the organisational culture, and it is the leaders of an organisation who determine how it functions, and it is their decision making which determines in particular, whether an organisation exhibits the practices and attitudes which go to make up a culture of

safety (Hopkins, 2005). The disaster at the Moura coal mine in central Queensland, which exploded in 1994, killing 11 men, presents an excellent illustration of the importance of safety culture in organisations. The accident inquiry revealed a culture, as a set of practices, focused on maximising production and largely oblivious to the potential for explosion (Hopkins, 1999). This accident is indicative of the systematic attention that was paid to production by managers at Moura and the systematic lack of attention paid to safety. This managerial focus shaped the whole culture of the mine. Organisational cultures may be detrimental to safety, not because leaders have chosen to sacrifice safety for the sake of production, but because they have not focused their attention on safety at all (Hopkins, 2005). Hopkins argues that if leaders attend to both production and safety, the organisations they lead will exhibit a culture which potentially emphasises both.

Pidgeon (1991) discusses a number of features that characterise a "good" safety culture: senior management commitment to safety; shared care and concern for hazards and a solicitude over their impacts upon people; Realistic and flexible norms and rules about hazards; and continual reflection upon practice through monitoring, analysis and feedback systems. Modern industrial organisations are facing strong pressures for change due to competition and change of generation (both technology and people) and at the same time they need to be able to ensure and demonstrate their reliability and safety in managing high-risk technological systems to the general public (Rieman & Oedewald, 2005). A central finding of the Columbia investigation report is the recommendation that NASA should address the "political, budgetary and policy decisions" that influenced the organisational structure, culture and systems safety which led to the flawed decision-making (CAIB, 2003). Leveson et al. (2004) propose a systems orientation approach that links system safety and engineering systems to address safety culture and other organisational dynamics in NASA.

7.3 Power and Politics in Organisations

Sagan's (1993) analysis of accidents and near-misses in the US nuclear weapons system provides compelling evidence that power and politics in complex organisations contribute to accidents, and furthermore emphasises the role of group interests in producing accident-prone systems.

Vaughn (1996) describes the Challenger accident as "social construction of reality" that allowed the banality of bureaucracy to create a habit of normalizing deviations from safe procedures. While Perrow (1999) concurs with Vaughan's account of the *Challenger* accident as an appropriately sociological and organisational explanation, he argues that Vaughan's analysis minimises the corruption of the safety culture, and more particularly drains this case of the extraordinary display of organisational power that overcame the objections of the engineers who opposed the launch. Perrow concludes that this was not the normalization of deviance or the banality of bureaucratic procedures and hierarchy or the product of an engineering "culture;" it was the exercise of organisational power.

Perrow (1994) discusses organisational politics where corporate leaders pay lip service to safety and use their power to impose risk on the many for the benefit of the few. He determines the reasons for corporate behaviour as: the latency period for a catastrophic accident to occur may be longer than any decision maker's career; few managers are punished for not putting safety first even after an accident, but will quickly be punished for not putting profits, market share or prestige first. Moreover, managers

may start to believe their own rhetoric about safety first because information that creates the awareness on lack of safety is suppressed for reasons of organisational politics. Sagan (1994) argues that even if organisational leaders place safety first and try to enforce this goal, clashes of power and interest at lower levels may defeat it.

It is essential to understand the role of politics and power in organisations as they have high potential to contribute to accident causation and disasters in complex sociotechnical systems. Sagan's (1993) study of nuclear weapons organisations found them to be infused with politics, with many conflicting interest at play both within the military command and control, and between military and civilian leaders. Sagan concludes that power and politics should be taken seriously and necessary not only to understand the organisational causes of accidents, but also to start the difficult process of designing reforms to enhance safety and reliability in organisations. Sagan encourages organisational theorists to study these organisational factors in order to bring the hazardous organisations' culture and operational practices to the public view.

8. Discussion and Conclusions

The underlying models of accidents can typically be grouped into three types (Hollnagel, 2004): sequential models, epidemiological models, and systemic models. The sequential and epidemiological models have contributed to the understanding of accidents; however, they are not suitable to capture the complexities and dynamics of modern sociotechnical systems. In contrast to these approaches, systemic models view accidents as emergent phenomena, which arise due to the complex and nonlinear interactions among system components. These interactions and events are hard to understand, and it is not sufficient to comprehend accident causation by employing the standard techniques in safety engineering alone, i.e. by analysing the failure modes of individual components using techniques such as Fault Tree Analysis and Failure Modes and Effects Analysis, or relating the accident to a single causal factor. Since the standard safety techniques concentrate on component failure, they cannot adequately capture the dysfunctional interactions between individual components operating without failure.

Accident models generally used for the prediction of accidents during the development of safety-critical system, in particular, are based on sequential models. Furthermore, traditional safety and risk analysis techniques such as Fault Tree Analysis and Probabilistic Safety Analysis are not adequate to account for the complexity of modern sociotechnical systems. The choice of accident model has consequence for how *post hoc* accident analysis and risk assessment is done, thus we need to consider the extension and development of systemic accident models both for accident analysis and for risk assessment and hazard analysis of complex critical systems.

Rasmussen's framework has been comprehensively and independently tested on the analysis of two Canadian public health disasters (Woo & Vicente, 2003) and on the Esso gas plant explosion accident in Australia (Hopkins, 2000). These case studies demonstrate the validity of Rasmussen's framework to explain the accident causation *a posteriori*. Further research is needed to extend this framework to predict accidents and to explore the applicability to risk and safety analysis of critical sociotechnical systems.

Similarly, STAMP has been applied to a number of case studies for *post hoc* accident analysis (e.g., Leveson et al., 2002; Johnson & Holloway, 2003b). There is a need for a methodology for the development of the STAMP model including guidelines for developing the control models and interpretation of the flawed control classification.

Some advances have been made in extending the STAMP model to conduct a proactive accident investigation in the early stages of system design. Leveson & Dulac (2005) discuss the use of STAMP model for hazard analysis, safety (risk) assessment, and as a basis for a comprehensive risk management system.

Cognitive systems engineering provides a framework for accident analysis in complex sociotechnical systems. A cognitive system is a system which can adapt its output to changes in the environment, with the purpose of staying in control of what the system does. A driver and his car is an example of a joint cognitive system which can adapt itself to changes in the environment, thereby keeping itself in control of its tasks (Huang, 2007). As contrasted with the contemporary techniques for human performance modelling, the joint cognitive systems paradigm stresses that modelling

the human operator as a separate system is not feasible; rather the human-machine is regarded as a whole where the dynamics and complexity of the interaction can be captured by providing a joint model (Hollnagel & Woods, 2005).

The recent advances in new systemic accident models, based on cognitive systems engineering, such as the Functional Resonance Accident Model (Hollnagel, 2004), should be investigated further and applied to the modelling of complex sociotechnical systems to understand the variability in human and system performance and how this relates to accident causation.

Although, formal methods have been applied successfully to the design and verification of safety-critical systems, they need to be extended to capture the many factors, including human behaviour and organisational aspects that are found in accidents and accident reports. Further research is needed to develop languages and semantics for modelling the various aspects of accidents in modern complex systems, such as: organisational, cultural and social properties, and human performance. However, formal methods have limitations in scalability to model complete sociotechnical systems, they need specialists in mathematics, and it should be noted that not every aspect of a complex system can be formalised in a mathematical sense. Why-Because Analysis is probably the most mature formal method for accident analysis. WBA has also been compared with other causal analysis methods; in particular the comparison with Rasmussen's AcciMap technique showed that the methodical approach employed by WBA produces greater precision in determining causal factors than does the informal approach of the AcciMap (Ladkin, 2005). However, a single case study is not sufficient to draw general results; comparisons of these methods need to be conducted on a large variety of sociotechnical systems in diverse domains.

Organisational sociologists have made significant contributions to the understanding of safety and accident causation in complex organisations managing and operating high-risk technological systems. There are two main schools of thought in sociology that have addressed the social, cultural and organisational aspects of safety and risk; they are identified as Normal Accident Theory (Perrow, 1984; Perrow, 1994) and High Reliability Organisation Theory (La Porte & Consolini, 1991; Roberts, 1989; Weick, 1987). These theories provide different explanations of safety and accident causation in complex organisations and offer alternative strategies for safety and risk management. The main premise of Normal Accident Theory is that even though risk prevention is taken seriously, there are several cognitive, social, cultural and system characteristics that over time accidents are inevitable. According to this theory, the organisations managing hazardous technologies exhibiting both high interactive complexity and tight coupling are candidates for accidents which cannot be avoided. High reliability theorists argue that accidents in the modern world can be prevented by complex organisations if appropriate organisational designs and management techniques are employed (Sagan, 1993). These theories generally emphasise the organisational aspect of accidents and tend to overlook the technical aspects, oversimplify the causes of accidents by focusing only on simple redundancy, and not considering accidents where component failure is not the cause (Marias et al., 2004).

System theoretical approach to safety provides a framework for modelling the technical, human, social and organisational factors in sociotechnical systems, including complex interactions among the system components. The sociotechnical system must

be treated as an integrated whole, and the emphasis should be on the simultaneous consideration of social and technical aspects of systems, including social structures and cultures, social interaction processes, and individual factors such as capability and motivation as well as engineering design and technical aspects of systems (Marias et al., 2004).

Future research is needed to comprehensively analyse the applicability of the new systemic models across a broader class of sociotechnical systems, particularly in the safety-critical sector such as patient safety, transportation, nuclear power, maritime, defence and aerospace. A number of studies have conducted comparisons of systemic accident models, particularly STAMP and Rasmussen's risk management framework (e.g. Johnson & de Almeida, 2007). Further studies should be conducted to compare and further develop the new systemic accident models in a variety of complex sociotechnical domains. Systems theory approaches to modelling and analysing safety are new, and there is a need to demonstrate that these models will be more effective than the traditional chain-of-event models (NAS, 2003).

Resilience Engineering is emerging as a new paradigm in safety management, where "success" is based on the ability of organisations, groups and individuals to anticipate the changing shape of risk before failures and harm occur (Hollnagel et al., 2006). Complex systems exhibit dynamic behaviour and continuously adapt their behaviour to account for the environmental disturbances. Such system adaptations cannot be pre-programmed during system design (Hollnagel, 2006b). According to Rasmussen's model (see Figure 9), a system may become unstable or lose control at the boundary of safety regulations. Thus resilience is the ability of organisations to maintain control in order to stay outside the accident region. Resilience engineering requires powerful methods, principles and tools that prevent this from taking place. Systemic accident models support the analytical aspects of resilience engineering, and STAMP has been applied to analyse the resilience of organisations confronted by high hazard and high performance demands (Leveson et al., 2006). For the predictive part, resilience engineering can be addressed, e.g., by means of a functional risk identification method, such as proposed by the Functional Resonance Accident Model (Hollnagel, 2004).

The complexity of modern sociotechnical systems poses a challenging area of interdisciplinary research in the development of new safety analysis and accident models involving researchers from engineering, social sciences, organisational theory, and cognitive psychology. Thus, there is a compelling need for researchers to step outside their traditional boundaries in order to capture the complexity of modern sociotechnical systems from a broad systemic view for understanding the multi-dimensional aspects of safety and modelling sociotechnical system accidents.

9. Acknowledgements

This research was initiated at the Defence Science and Technology Organisation, under the Critical Systems Development Task JTW 04/061, sponsored by the Defence Materiel Organisation. I am particularly indebted to Dr. Tony Cant, DSTO for his continuous encouragement and inspiring discussions on safety-critical systems. I am grateful to Drs. Brendan Mahony and Jim McCarthy, High Assurance Systems Cell, Command, Control, Communications and Intelligence Division, DSTO who have assisted me greatly with their expertise in many "formal" aspects of safety-critical systems and accident modelling research, and for the many stimulating afternoon discussions.

I would like to thank Professor Stephen Cook and Associate Professor David Cropley, Director and Deputy-Director respectively, of the Defence and Systems Institute at the University of South Australia for their support in the writing of this report and in providing a congenial atmosphere for system safety research.

I am grateful to many members of the Australian Safety Critical Systems Association for their constructive comments and to the anonymous reviewers for their feedback on a version of this research work that was presented at the 12th Australian Workshop on Safety Related Programmable Systems (SCS'07). Finally, I would like to express my appreciation to Dr. Michelle Grech, Human Factors Group, Maritime Platforms Division, DSTO for reviewing this report and for the many helpful comments that greatly assisted in the preparation of the final version.

10. References

AAIB (1994). *U.S. Army Black Hawk Helicopters 87-26000 and 88-26060: Volume 1*. Executive Summary: UH-60 Black Hawk Helicopter Accident, 14 April, USAF Aircraft Accident Investigation Board.

Anderson, S. O., Bloomfield, R. E., & Cleland, G. L. (1998a). *Guidance on the use of Formal Methods in the Development and Assurance of High Integrity Industrial Computer System, Parts I and II*. Working Paper 4001, European Workshop on Industrial Computer Systems (EWICS) Technical Committee 7.
http://www.ewics.org/docs/formal-methods-subgroup

Anderson, S. O., Bloomfield, R. E., and Cleland, G. L. (1998b). *Guidance on the use of Formal Methods in the Development and Assurance of High Integrity Industrial Computer Systems, Parts III A Directory of Formal Methods*. Working Paper 4002, European Workshop on Industrial Computer Systems (EWICS) Technical Committee 7.
http://www.ewics.org/docs/formal-methods-subgroup

ATEA (1998). *Def (Aust) 5679: The Procurement of Computer-Based Safety-Critical Systems*. Australian Defence Standard, August, Australia: Army Technology Engineering Agency.

Benner, L. (1975). Accident investigation: Multilinear events sequencing methods. *Journal of Safety Research*, 7(2), 67-73.

Besnard, D. & Baxter, G. (2003). *Human compensations for undependable systems*. Technical Report Series, CS-TR-819, Newcastle upon Tyne: University of Newcastle upon Tyne. http://www.dirc.org.uk/publications/techreports/papers/12.pdf

Bjørner, D. & Druffel, L. (1990). Position statement: ICSE-12 Workshop on Industrial Experience Using Formal Methods. *Proceedings of the 12th International Conference on Software Engineering*, 264-266, March 26-30, Nice, France.

Blackman, H., Gertman, D. & Hallbert, B. (2000). The need for organisational analysis. *Cognition, Technology & Work*, 2, 206-208.

Booch, G. (1994). *Object-Oriented Analysis and Design with Applications.* 2nd Ed., Menlo Park, CA: Addison-Wesley.

Bowen, J. & Stavridou, V. (1993). Safety Critical Systems: formal methods and standards. *Software Engineering Journal*, 8(4), 189-209, UK: IEE.

Buede, D. M. (2000). *The Engineering Design of Systems: Models and Methods*. New York: Wiley.

Burns, C. P. (2000). *Analysing Accident Reports Using Structured and Formal Methods*. Ph.D. Thesis, February, Glasgow: The University of Glasgow.
http://www.dcs.gla.ac.uk/research/gaag/colin/thesis.pdf

Burns, C. P., Johnson, C. W. & Thomas, M. (1997). *Agents and actions: Structuring*

human factors accounts of major accidents. Technical Report TR-1997-32, Department of Computing Science, October, Glasgow: University of Glasgow.

Butler, R. W., Carreño, V. A., Di Vito, B. L., Holloway, C. M. & Miner, P. S. (2002). *NASA Langley's Research and Technology-Transfer Program in Formal Methods*. Assessment Technology Branch, Hampton, Virginia: NASA Langley Research Center. http://shemesh.larc.nasa.gov/fm/NASA-over.pdf

CAIB (2003). *Columbia Accident Investigation Board Report Volume I*. Washington, D.C.: Columbia Accident Investigation Board.

CEL (2007). Risk Management Framework, Cognitive Engineering Laboratory, University of Toronto, http://www.mie.utoronto.ca/labs/cel/research/frameworks/risk.htm (1992-2007)

Collins, J., Hall, N. &. Paul, L. A. (2004). Counterfactuals and Causation: History, Problems, and Prospects, Chapter 1, In Collins, J., Hall, N. &. Paul, L. A. (Eds.), *Causation and Counterfactuals*. Cambridge, MA: The MIT Press.

Clarke, E. M. & Wing, J. M. (1996). *Formal Methods: State of the Art and Future Directions*. Report CMU-CS-96-178, School of Computer Science, Pittsburgh PA: Carnegie Mellon University.

Clarkson, J., Hopkins, A. & Taylor, K. (2001): *Report of the Board of Inquiry into F-111 (Fuel Tank) Deseal/Reseal and Spray Seal Programs - Vol. 1*. Canberra, ACT: Royal Australian Air Force. http://www.defence.gov.au/raaf/organisation/info_on/units/f111/Volume1.htm

Dien, Y., Llory, M. & Montmayeul, R. (2004). Organisational accidents investigation methodology and lessons learned. *Journal of Hazardous Materials*, 111, 147-153.

Eiter, T. & Lukasiewicz, T. (2002). *Complexity Results for Explanations in The Structural-Model Approach*. INFSYS Research Report 1843-01-08, July, Institut für Informationssysteme, Abtg. Wissensbasierte Systeme Technische, Universität Wien Favoritenstraße 9-11 A-1040, Wien, Austria.

Ferry, T. S. (1988). *Modern Accident Investigation and Analysis*. Second Edition, New York: J. Wiley.

Fields, R. E., Wright, P. C. & Harrison, M. D. (1995). A task centred approach to analysing human error tolerance requirements. In *RE'95 - Second IEEE International Symposium on Requirements Engineering*, York.

FME (2004). *Formal Methods Europe*. http://www.fmeurope.org (accessed: 21 December 2006).

FMVL (2007). Formal Methods Virtual Library. http://vl.fmnet.info/ (last updated: 24 August 2007).

GAAG (2006). Glasgow Accident Analysis Group, Dept. of Computer Science, University of Glasgow, Scotland. http://www.dcs.gla.ac.uk/research/gaag/

GAO (1997). *Operation Provide Comfort: Review of Air Force Investigation of Black Hawk Fracticide Incident*. GAO/OSI-9804, Office of Special Investigations, Washington DC: US General Accounting Office.

Gaudel, Marie-Claude (1994). Formal specification techniques. *Proceedings of the 16th International Conference on Software Engineering*. Sorrento, Italy, 223-227, Los Alamitos, CA: IEEE Computer Society Press.

Gomaa, H. (1986). Software development for real-time systems, *Communications of the ACM*, 29(7), 657-668.

Harel, D. (1987). Statecharts: A Visual Formalism for Complex Systems. *Science of Computer Programming*, 8, 231-273.

Haveland, K. & Lowry, M. (2001). Formal Analysis of a Space-Craft Controller using SPIN. *IEEE Transactions on Software Engineering*, 27(8), 749-765.

Hayhurst, K. J. & Holloway, C. M. (2003). *Second Workshop on the Investigation and Reporting of Incidents and Accidents*. IRIA 2003, National Aeronautics and Space Administration Washington, DC 20546-0001, NASA/CP-2003-212642, Hampton, VA: NASA Langley Research Center.
http://techreports.larc.nasa.gov/ltrs/ or http://ntrs.nasa.gov

Heinrich, H. W., Petersen, D. & Roos, N. (1980). *Industrial Accident Prevention*. New York: McGraw-Hill.

Hinchey M. G. & Bowen, J. P. (Eds.). (1995). *Applications of Formal Methods*. International Series in Computer Science, UK: Prentice Hall.

Hitchcock, C. (2002). Probabilistic Causation. In Edward N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy (Fall 2002 Edition)*.
http://plato.stanford.edu/archives/fall2002/entries/causation-probabilistic/

Hoare, C. A. R. (1985). *Communicating Sequential Processes*. Hemel Hempstead: Prentice Hall.

Höhl, M. & Ladkin, P. (1997). *Analysing the 1993 Warsaw Accident with a WB-Graph*. Article RVS-Occ-97-09, 8 September, Faculty of Technology, Bielefeld University. http://www.rvs.uni-bielefeld.de

Hollnagel, E. (1998). *Cognitive Reliability and Error Analysis Method*. Oxford: Elsevier Science.

Hollnagel, E. (2001). Anticipating Failures: What should Predictions be About? In *The Human Factor in System Reliability – Is Human Performance Predictable?* RTO Meeting Proceedings 32, RTO-MP-32, January, Research and Technology Organization, North Atlantic Treaty Organization, Cedex, France: RTO/NATO.

Hollnagel, E. (2002). Understanding Accidents – From Root Causes to Performance Variability. *IEEE 7th Human Factors Meeting*, Scottsdale, Arizona.

Hollnagel, E. (2004). *Barriers and Accident Prevention*. Hampshire: Ashgate.

Hollnagel, E. (2006a). *CREAM - Cognitive Reliability and Error Analysis Method*. http://www.ida.liu.se/~eriho/CREAM_M.htm

Hollnagel, E. (2006b). Resilience – the Challenge of the Unstable. In Hollnagel, E., Woods, D. D. & Leveson, N. (2006). *Resilience Engineering: Concepts and Precepts*. Aldershot: Ashgate.

Hollnagel, E. & Woods, D. D. (1983). Cognitive Systems Engineering: New wine in new bottles. *International Journal of Man-Machine Studies*, 18, 583-600. Reprinted in International Journal of Human-Computer Studies, 1999, 51, 339-356.

Hollnagel, E. & Woods, D. (2005*). Joint Cognitive Systems: Foundations of Cognitive Systems Engineering*. New York: Taylor & Francis.

Hollnagel, E., Woods, D. D. & Leveson, N. (2006). *Resilience Engineering: Concepts and Precepts*. Aldershot: Ashgate.

Hopkins, A. (1999*). Managing Major Hazards: The Lessons of the Moura Mine Disaster*. Sydney: Allen and Unwin.

Hopkins, A. (2000). *Lessons from Longford: The Esso Gas Plant Explosion*. Sydney: CCH.

Hopkins, A. (2005). *Safety, Culture and Risk: The Organisational Causes of Disasters*. Sydney: CCH.

Huang, Yu-Hsing (2007). *Having a New Pair of Glasses Applying Systemic Accident Models on Road Safety*. Dissertation No. 1051, Department of Computer and Information Science, Linköping, Sweden: Linköping University.

IEC 61508 (1998-2000). *Functional safety of electrical/electronic/programmable electronic safety-related system*. Parts 1 to 7, Geneva, Switzerland: International Electrotechnical Commission.

Johnson, C. W. (2000). *Models and use of Counter-Factual reasoning in Accident Investigations*. Department of Computing Science, University of Glasgow, Scotland. http://www.dcs.gla.ac.uk/~johnson/

Johnson, C. W. (2003). *The Failure of Safety-Critical Systems: A Handbook of Accident and Incident Reporting*. October, Glasgow, Scotland: Glasgow University Press. http://www.dcs.gla.ac.uk/~johnson/book

Johnson, C. J. & Botting, R. M. (1999). Using Reason's model of Organisational Accidents in Formalising Accident Reports. *Cognition, Technology & Work*, 1, 107-118.

Johnson, C. W. & de Almeida, I. M. (2007). An investigation into the loss of the Brazilian space programme's launch vehicle VLS-1 V03. *Safety Science*, In Press, doi:10.1016/j.ssci.2006.05.007.

Johnson, C. & Holloway, C. M. (2003a). A Survey of Logic Formalisms to Support Mishap Analysis. *Reliability Engineering & System Safety*, 80(3), 271-291.

Johnson, C., & Holloway, C. M. (2003b). The ESA/NASA SOHO Mission Interruption: Using the STAMP Accident Analysis Technique for a Software Related `Mishap'. *Software: Practice and Experience*, 33(12), 1177-1198.

Johnson, C. W. & Holloway, C. M. (2007). A Longitudinal Analysis of the Causal Factors in Major Maritime Accidents in the USA and Canada (1996-2006). *Proceedings of the 15th Safety-Critical Systems Symposium*, Bristol, UK, 13-15 February, Redmill, F. & Anderson, T. (Eds.), The Safety of Systems, 85-94, Springer.

Khosla, S. (1988). *System Specification: A Deontic Approach.* PhD thesis, Imperial College of Science and Technology, UK: University of London.

Kroes, P., Franssen, M., van de Poel, Ibo. & Ottens, M. (2006). Treating socio-technical systems as engineering systems: some conceptual problems. *Systems Research and Behavioral Science*, 23(6), 803-814.

La Porte, T. R. & Consolini, P. M. (1991). Working in Practice but Not in Theory: Theoretical Challenges of "High-Reliability Organizations". *Journal of Public Administration Research and Theory*, 1(1), 19-47.

Ladkin, P. B. (1999). *A Quick Introduction to Why-Because Analysis*, 1 March, Faculty of Technology, Bielefeld University. http://www.rvs.uni-bielefeld.de

Ladkin, P.B. (2005). *Why-Because Analysis of the Glenbrook, NSW Rail Accident and Comparison with Hopkins's Accimap*. Report RVS-RR-05-05, 19 December, Faculty of Technology, Bielefeld University. http://www.rvs.uni-bielefeld.de

Ladkin, P. B. & Loer, K. (1998). *Why-because analysis: Formal reasoning about incidents*. Technical Report RVS-Bk-98-01, Faculty of Technology, Bielefeld University http://www.rvs.uni-bielefeld.de

Ladkin, P. B. & Stuphorn, J. (2003). Two Causal Analyses of the Black Hawk Shootdown During Operation Provide Comfort, *Proceedings of the 8th Australian Workshop on Safety Critical Software and Systems*, Peter Lindsay and Tony Cant (Eds.), Conferences in Research and Practice in Information Technology, Volume 33, Canberra: Australian Computer Society.

Lamport, L. (1994). The temporal logic of actions. ACM Transactions on Programming Languages and Systems (TOPLAS), 16(3), 872-923.

Lamsweerde, A. V. (2000). Formal Specification: A Roadmap. *Proceedings of the Conference on The Future of Software Engineering*, Limerick, Ireland, 147-159, ACM Press.

LaRC (2004). The CAUSE Project, Research on Accident Analysis, NASA Langley Formal Methods Site. http://shemesh.larc.nasa.gov/fm/fm-now-cause.html

Leveson, N. G. (1986). Software Safety: Why, What, and How, *Computing Surveys*, 18, 2 June.

Leveson, N. (1993). An Investigation of the Therac-25 accidents. *IEEE Computer*, 26, 18-41.

Leveson, N. G. (1995). *Safeware: System Safety and Computers*. Reading. MA: Addison-Wesley.

Leveson, N. (2001). *Evaluating Accident Models using Recent Aerospace Accident -. Part I: Event–based Models*. Technical Report, Aeronautics and Astronautics Department, Massachusetts Institute of Technology, June 28, Cambridge, MA: MIT.

Leveson, N. G. (2002). *System Safety Engineering: Back to the Future*. Aeronautics and Astronautics Department, Massachusetts Institute of Technology, Cambridge, MA: MIT. http://sunnyday.mit.edu/book2.pdf

Leveson, N. (2003). *A New Approach to System Safety Engineering*, Aeronautics and Astronautics Department, Massachusetts Institute of Technology, Cambridge, MA: MIT. http://sunnyday.mit.edu

Leveson, N. (2004). A New Accident Model for Engineering Safer Systems, *Safety Science*, 42, 237-270.

Leveson, N. G., Allen, P. & Storey, Margaret-Anne. (2002). The Analysis of a Friendly Fire Accident using a Systems Model of Accidents. *Proceedings of the 20th International System Safety Conference*, Denver, Colorado, 5-9 August.

Leveson, N., Cucher-Gershenfeld, J., Barrett, B., Brown, A., Carroll, J., Dulac, N., Fraile, L. & Marais, K. (2004). *Effectively addressing NASA's organization and safety culture*. Engineering Systems Division Symposium, Massachusetts Institute of Technology, March 29-31, Cambridge, MA: MIT.

Leveson, N. G. & Dulac, N. (2005). Safety and Risk-Driven Design in Complex Systems-of-Systems. *1st NASA/AIAA Space Exploration Conference*, Orlando.

Leveson, N., Dulac, N., Zipkin, D., Cutcher-Gershenfeld, J., Carroll, J. & Barrett, B (2006). Engineering Resilience into Safety-Critical Systems. In Hollnagel, E., Woods, D. D. & Leveson, N. (2006). *Resilience Engineering: Concepts and Precepts*. Aldershot: Ashgate.

Leveson, N. G., Heimdahl, M .P. E., Hildreth, H. & Reese, J. D. (1994). Requirements specification for process-control systems, *IEEE Transactions on Software Engineering*, 20(9), 684-707, September.

Lewis, D. (1973). Causation. *Journal of Philosophy*, 70 (17), 556-567.

Lindsay, P. (1998). A Tutorial Introduction to Formal Methods. *Proceedings 3rd Australian Workshop on Industrial Experience with Safety Critical Systems and Software*, 13, 29-37, November, Canberra: Australian Computer Society.

Marais, K., Dulac, N., & Leveson, N. (2004). *Beyond Normal Accidents and High Reliability Organizations: The Need for an Alternative Approach to Safety in Complex Systems*, ESD Symposium, Massachusetts Institute of Technology, Cambridge, MA: MIT.

Maurino, D., Reason, J. T., Johnston, N. & Lee, R. (1995). *Beyond aviation human factors*, Aldershot: Ashgate.

Menzies, P. (2001), Counterfactual Theories of Causation, In Edward N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy (Spring 2001 Edition)*. http://plato.stanford.edu/archives/spr2001/entries/causation-counterfactual/

NAS (2003). *Securing the Future of U.S. Air Transportation: A System in Peril*, Committee on Aeronautics Research and Technology for Vision 2050, Aeronautics and Space Engineering Board, Washington, DC: National Academy of Sciences.

Parasuraman, R. (1997). Humans and Automation: use, misuse, disuse, abuse. *Human Factors*, 39(2), 230-253.

Pearl, J. (2000). *Causality: Models, Reasoning, and Inference*, UK: Cambridge University Press.

Perrow, C. (1984). *Normal Accidents: Living with High-Risk Technologies*. New York: Basic Books.

Perrow, C. (1994). The Limits of Safety: The Enhancements of a Theory of Accidents. *Journal of Contingencies & Crisis Management*, 2, 4, 212-220.

Perrow, C. (1999). Y2K as a normal accident. *International Conference on Disaster Management and Medical Relief*, June 14-16, Amsterdam.

Pidgeon, N. (1991). Safety Culture and Risk Management in Organizations. *Journal of Cross-Cultural Psychology*, 22(1), 129-140.

Rasmussen, J. (1997). Risk Management in a Dynamic Society: A Modelling Problem, *Safety Science*, 27, 2, 183-213.

Rasmussen, J. & Svedung. X. (2000). *Proactive Risk Management in a Dynamic Society*. Karl, Sweden: Swedish Rescue Services Agency.

Reason, J. (1990). *Human Error*. Cambridge: Cambridge University Press.

Reason, J. (1997). *Managing the Risks of Organizational Accidents*. England: Ashgate.

Reason, J. (2000). Human error: models and management. *British Medical Journal*, 320, 768-770.

Reason, J. T., Carthey, J. & de Laval, M. R. (2000). Diagnosing 'Vulnerable System Syndrome': a prerequisite to effective risk management. *Quality Health Care*, 10, 21-25.

Rieman, T. & Oedewald, P. (2005). Cultural approach to organisations and management of change. In Nuutinen, M. & Luoma, J. (Eds.), *Human practice in the life cycle of complex systems: challenges and methods*. VTT Publications 582, Helsinki: VTT Technical Research Centre of Finland.

Roberts, K. H. (1989). New Challenges in Organization Research: High Reliability

Organizations. *Industrial Crisis Quarterly*, 3(2), 111-125.

Rogers Commission Report (1986). *Report of the Presidential Commission on the Space Shuttle Challenger Accident*. June 6, Washington, D.C.: NASA. http://history.nasa.gov/rogersrep/genindex.htm

Sagan, S. (1993). *Limits of Safety: Organizations, Accidents, and Nuclear Weapons*. Princeton, New Jersey: Princeton University Press.

Sagan, S. (1994). Toward a Political Theory of Organizational Reliability. *Journal of Contingencies & Crisis Management*, 2, 4, 228-240.

Shappell, S. A. & Wiegmann, D. A. (2000). *Human factors analysis and classification system - HFACS*. Report DOT/FAA/AM-00/7, Washington, DC: Department of Transportation, FAA.

Shorrock, S., Young, M. & Faulkner, J. (2003). Who moved my (Swiss) cheese? *Aircraft and Aerospace*, January/February, 31-33.

Skelt, S. (2002). *Methods for accident analysis*. Report No. ROSS (NTNU) 2000208, Norwegian University of Science and Technology, Trondhein: NTNU.

Snook, S. (2002). *Friendly Fire: The Accidental Shootdown of U.S. Black Hawks over Northern Iraq*, Princeton, New Jersey: Princeton University Press.

Shrivastava, P. (1992). *Bhopal: Anatomy of a Crisis*. Second Edition, London: Paul Chapman.

Svedung, X. & Rasmussen, J. (2002). Graphic representation of accident scenarios: mapping system structure and the causation of accidents. *Safety Science*, 40, 397-417.

Thomas, M. (1994). A proof of incorrectness using the LP theorem prover: The editing problem in the Therac-25. *High Integrity Systems*, 1(1), 35-49.

Trist, E. L. & Bamforth, K. W. (1951). Some Social and Psychological Consequences of the Longwall Method of Coal-Getting. *Human Relations*, 4, 3-39.

Vaughn, D. (1996). *The Challenger Launch Decision: Risky Technology, Culture and Deviance at NASA*. Chicago: University of Chicago Press.

Vernez, D., Buchs, D. & Pierrehumbert, G. (2003). Perspectives in the use of coloured Petri nets for risk analysis and accident modelling. *Safety Science*, 41, 445-463.

Vicente, K. J. (1999). *Cognitive Work Analysis: Towards Safe, Productive, and Healthy Computer-Based Work*. Mahwah, NJ: Lawrence Erlbaum.

Vicente, K. J., Roth, E. M. & Mumaw, R. J. (2001). How do operators monitor a complex, dynamic work domain? The impact of control room technology. *International Journal of Human-Computer Studies*, 54, 831-856.

Vicente, K. J., Mumaw, R. J. & Roth, E. M. (2004). Operator monitoring in a complex dynamic work environment: A qualitative cognitive model based on field observations. *Theoretical Issues in Ergonomics Science*, 5(5), 359-384.

Wagenaar, W. A., Groeneweg, J., Hudson, P. T. W. & Reason J. T. (1994). Safety in the oil industry. *Ergonomics*, 37, 1999-2013.

Weick, K. E. (1987). Organizational Culture as a Source of High Reliability. *California Management Review*, 29(2), 112-127.

Wieringa, R. J. & Meyer, J. -J. Ch. (1994). Applications of Deontic Logic in Computer Science: A Concise Overview, In J.-J. Ch. Meyer & R.J. Wieringa (Eds.), *Deontic Logic in Computer Science: Normative System Specification*. Chichester: John Wiley and Sons.

Wildman, L. (2002). Requirements reformulation using formal specification : A case study, In Lakos, C., Esser, R., Kristensen, L. M., and Billington, J. (Eds.), *Proceedings of Workshop on Formal Methods Applied to Defence Systems*, June, Vol. 12 of Conferences in Research and Practice in Information Technology, 75-83, Canberra: Australian Computer Society..

Wing, J. M. (1990). A specifier's introduction to formal methods**.** *Computer*, 23(9), 8-24, September, New Jersey: IEEE.

Woo, D. M. & Vicente, K. J. (2003). Sociotechnical systems, risk management, and public health: comparing the North Battleford and Walkerton outbreaks. *Reliability Engineering & System Safety*, 80, 253-269.

Woods, D. D., Johannesen, L. J. & Sarter, N. B. (1994*). Behind Human Error: Cognitive Systems, Computers and Hindsight*. SOAR Report 94-01, Wright-Patterson Air Force Base, Ohio: CSERIAC.

Yourdon, E. (1989). Modern Structured Analysis. Englewood Cliffs, NJ: Yourdon Press.

DEFENCE SCIENCE AND TECHNOLOGY ORGANISATION DOCUMENT CONTROL DATA	1. PRIVACY MARKING/CAVEAT (OF DOCUMENT)

2. TITLE	3. SECURITY CLASSIFICATION (FOR UNCLASSIFIED REPORTS THAT ARE LIMITED RELEASE USE (L) NEXT TO DOCUMENT CLASSIFICATION)
A Review of Accident Modelling Approaches for Complex Critical Sociotechnical Systems	Document (U) Title (U) Abstract (U)
4. AUTHOR(S)	5. CORPORATE AUTHOR
Zahid H. Qureshi	DSTO Defence Science and Technology Organisation PO Box 1500 Edinburgh South Australia 5111 Australia

6a. DSTO NUMBER	6b. AR NUMBER	6c. TYPE OF REPORT	7. DOCUMENT DATE
DSTO-TR-2094	AR 014-089	Technical Report	January 2008

8. FILE NUMBER	9. TASK NUMBER	10. TASK SPONSOR	11. NO. OF PAGES	12. NO. OF REFERENCES
	C3ID DMO 07-007	DMO	66	124

13. URL on the World Wide Web	14. RELEASE AUTHORITY
http://www.dsto.defence.gov.au/corporate/reports/DSTO-TR- 2094.pdf	Chief, Command, Control, Communications and Intelligence Division

15. SECONDARY RELEASE STATEMENT OF THIS DOCUMENT

Approved for public release

16. DELIBERATE ANNOUNCEMENT

No Limitations

17. CITATION IN OTHER DOCUMENTS

18. DSTO RESEARCH LIBRARY THESAURUS http://web-vic.dsto.defence.gov.au/workareas/library/resources/dsto_thesaurus.htm

Accidents, Sociotechnical systems, Safety-Critical Systems

19. ABSTRACT

The increasing complexity in highly technological systems such as aviation, maritime, air traffic control, telecommunications, nuclear power plants, defence and aerospace, chemical and petroleum industry, and healthcare and patient safety is leading to potentially disastrous failure modes and new kinds of safety issues. Traditional accident modelling approaches are not adequate to analyse accidents that occur in modern sociotechnical systems, where accident causation is not the result of an individual component failure or human error. This report provides a review of key traditional accident modelling approaches and their limitations, and describes new system-theoretic approaches to the modelling and analysis of accidents in safety-critical systems. It also discusses current research on the application of formal (mathematically-based) methods to accident modelling and organisational theories on safety and accident causation. This report recommends new approaches to the modelling and analysis of complex systems that are based on systems theory and interdisciplinary research, in order to capture the complexity of modern sociotechnical systems from a broad systemic view for understanding the multi-dimensional aspects of safety and accident causation.